

ON F AND E , IN DFT

PAUL E. LAMMERT

*Dept. of Physics, 104B Davey Lab
 Pennsylvania State University
 University Park, PA 16802-6300*

ABSTRACT. Rigorous mathematical foundations of density functional theory are revisited, with some use of infinitesimal (nonstandard) methods. A thorough treatment is given of basic properties of internal energy and ground-state energy functionals along with several improvements and clarifications of known results. A simple metrizable topology is constructed on the space of densities using a hierarchy of spatial partitions. This topology is very weak, but supplemented by control of internal energy, it is, in a rough sense, essentially as strong as L^1 . Consequently, the internal energy functional F is lower semicontinuous with respect to it. With separation of positive and negative parts of external potentials, very badly behaved, even infinite, positive parts can be handled. Confining potentials are thereby incorporated directly into the density functional framework.

CONTENTS

1. Introduction	2
1.1. Motivations	2
1.2. What is here, old and new	2
1.3. Particle types	4
2. Internal energy as function of state (\mathcal{E})	4
2.1. Kinetic energy	4
2.2. Interaction energy	4
2.3. $\mathcal{E} : \mathcal{H} \rightarrow \overline{\mathbb{R}}$ is lower semicontinuous	5
2.4. Extension to mixed states	6
3. Density (ρ)	7
3.1. dens is continuous	7
3.2. Tameness and nearstandardness	7
4. Internal energy as function of density (F° and F)	9
4.1. Constrained energy minimization	9
4.2. Sobolev norms and kinetic energy bounds	10
4.3. Effective domain	11
4.4. $\mathcal{J}_{\mathcal{N}} \subset L^1 \cap L^3 \cap W^{1,1}$	13
4.5. Internal energy is lower semicontinuous on $\text{Dens}_{\mathcal{N}}$	13
5. Weak- \mathfrak{P} topology on $\text{Dens}_{\mathcal{N}}$	14
5.1. Hierarchy of partitions	14
5.2. How weak is it?	15
5.3. Approximation by projection	15
5.4. Weak- \mathfrak{P} with internal energy control is as good as L^1	16
6. Ground-state energy, or ‘energy floor’ (E)	17
6.1. Bounded external potentials	17

E-mail address: pel1@psu.edu.

Date: February 7, 2014.

6.2.	$[0, +\infty]$ -valued external potentials	19
6.3.	F -norm, and relatively F -small perturbations	19
6.4.	$L^{3/2} + L^\infty$	20
6.5.	Concavity and continuity	20
	Appendix A. Selective functional analysis refresher	21
A.1.	Lebesgue (L^p) and Sobolev ($H^1, W^{1,1}$) spaces	21
A.2.	the concept of ‘topology’	22
	Appendix B. Infinitesimals	22
B.1.	seriously	22
B.2.	base datatypes	23
B.3.	limits and continuity	23
B.4.	metric, function, and topological spaces	24
B.5.	*-transform and Leibniz’ Principle	24
B.6.	overspill	25
	References	25

1. INTRODUCTION

1.1. Motivations. In the density functional theory (DFT) literature, whenever the matter of mathematically rigorous foundations arises, a 1983 paper by Elliott Lieb[1] justly looms large. It propounds what could reasonably be called the “standard framework”. Although (or perhaps because) the community seems generally to regard it as a satisfactory foundation, the literature gives evidence that it is not widely understood. Standard DFT textbooks[2, 3, 4, 5] and reviews[6] spill very little ink on such matters since they have much else to cover. What few expository treatments exist[7, 4] hew very close to Lieb, and are not easy to obtain. Consequently, one more such exposition may have a place. However, while these notes are a fairly complete treatment of the most fundamental matters, it does not just add another set of footprints to the same patch of ground. The standard framework locates densities in $L^1(\mathbb{R}^3) \cap L^3(\mathbb{R}^3)$ and external one-body potentials in the dual space $L^\infty(\mathbb{R}^3) + L^{3/2}(\mathbb{R}^3)$, with the internal energy function F on the former and the ground-state energy E on the latter in a relation of Fenchel conjugacy. Rigid adherence to this framework is both unnecessarily limiting and physically distorting. Consider first the density side. In its primitive conception, density is a *measure*, telling us the mass in every region. The natural topologies for measures are weak topologies. This is reinforced by the fact (§4.5.3, apparently not noticed before, that F is lower semicontinuous with respect to topologies weaker than weak- L^1 . Keeping in mind that densities are non-negative and normalized, not general elements of a *vector* space is important to seeing that. On the potential side, the standard framework is even more constricting. It is good physics, and may even be useful, to allow potentials to have vastly more ill-behaved positive parts than negative parts. Why is $E(v) > -\infty$ for every $v \in L^\infty + L^{3/2}$? The general theory of Fenchel conjugacy, by itself, certainly does not make that expected. A very clear understanding (Thm. 6.2) is reached by maintaining a clear view of what is required to keep positive or negative unboundedness under control.

1.2. What is here, old and new. This section gives a brief tour of the contents of the paper, emphasizing the new results and putting it in context with some initial casual motivation.

It all begins with the following problem of conventional Hilbert-space quantum mechanics. A system of \mathcal{N} indistinguishable particles (electrons, if you wish) with kinetic energy (operator) T and mutual interaction W is placed in an external one-body potential v , and we want to find the ground state and its energy. That is, a minimizer of $\langle \psi | T + W + \underline{v} | \psi \rangle$ is sought:

$$E(v) = \inf_{\psi} \{ \langle \psi | T + W | \psi \rangle + \langle \psi | \underline{v} | \psi \rangle \}.$$

\underline{v} here is the \mathcal{N} -body potential obtained by adding a copy of v for each particle in the system. But building \underline{v} is a step in the wrong direction. The external potential energy depends on the state only

through its one-particle density. So, instead of minimizing over all states in one go, why not split it into two stages? First, for each density ρ , find the state with lowest internal energy,

$$F^\circ(\rho) = \inf_{\text{dens } \psi=\rho} \langle \psi | T + W | \psi \rangle. \quad (1.1)$$

Since all states with density ρ have the same external energy, only the best among them has any chance of being a ground state. Then, in stage two, minimize over densities:

$$E(v) = \inf_{\rho} \left\{ F^\circ(\rho) + \int v \rho \right\}. \quad (1.2)$$

This is the constrained-search approach of Levy[8] and Lieb[1]. The original problem involved particular \mathcal{N} , W and v . This final formula suggests an entirely different viewpoint. The *system* specified by \mathcal{N} and W is fixed, and we study its *response* to *many different* external potentials. All the relevant properties of the system are encoded in $F^\circ(\rho)$.

This may remind the reader of statistical mechanics. In a prototypical magnetic system, there are an enormous number of internal degrees of freedom, but the only thing relevant for the interaction with an externally imposed magnetic field is the net magnetization. The free energy as a function of magnetization is analogous to $F^\circ(\rho)$, and the free energy as function of external field, to $E(v)$. The two are related by a Legendre transformation, similarly to (1.2). But thermodynamic conjugacy works both ways. Do we also have

$$F^\circ(\rho) = \sup_v \left\{ E(v) - \int v \rho \right\} ? \quad (1.3)$$

Questions now come hard and fast; they are the subject of this paper. We have lots of infima here. Are there, in fact, minimizers? Consider the functions $f(x) = 1/x$ and $g(x) = (\text{ if } x = 0 \text{ then } 1 \text{ else } x^2)$. Then, $\inf_x f(x) = \inf_x g(x) = 0$, but neither function actually takes the value 0 anywhere. Take minimizing sequences $f(x_n) \xrightarrow{n \rightarrow \infty} 0$, and $g(y_n) \xrightarrow{n \rightarrow \infty} 0$. We would hope that at least subsequences converge to minimizers. But it does not work. There is no subsequence of (x_n) that converges at all. The sequence (y_n) converges to zero, but the function g suddenly jumps up. In the case of f , there is a failure of *compactness*, in the case of g , of *lower-semicontinuity*. (If g suddenly jumped *down* at zero, all would be well, so only “half” of continuity is really needed.) In the context of function spaces there are many distinct, legitimate and useful concepts of convergence. When we say, “ F is lower semicontinuous”, we want to make the strongest possible statement, so we seek the *weakest topology* with respect to which it is true. Lower-semicontinuity is a pervasive theme in the following pages.

In §2, we study the internal energy as a function of pure and mixed states, and its lower-semicontinuity. §3 turns attention to density, where the important results involve continuity of the map from states to densities and compactness of the set of low-energy states with given density. Pure state (F°) and mixed-state (F) internal energy as a function of density is studied in §4. Several results here are new. Lower-semicontinuity with respect to topologies much weaker than the L^1 topology is demonstrated. A general method of transporting densities and wavefunctions is used to make the well-known demonstration of an upper bound on F° in terms of the H^1 Sobolev norm of $\sqrt{\rho}$ less *ad hoc* and to facilitate the calculation of a kinetic energy bound. We point out that not only do finite-energy densities satisfy $\sqrt{\rho} \in H^1(\mathbb{R}^3)$ as is well-known, but also $\rho \in W^{1,1}(\mathbb{R}^3)$. Section 5 presents the construction of a convenient very weak yet metrizable topology on the space of densities, based on a hierarchy of partitions of space into cubical lattices. F is lower semicontinuous with respect to this weak- \mathfrak{P} topology. Despite its weakness, the weak- \mathfrak{P} topology is, roughly speaking, essentially as strong as L^1 under conditions of control on the internal energy. Finally, in §6, attention is turned to the ground-state energy $E(v)$. Here we make a very strong, and uncustomary, distinction between unbounded positive and unbounded negative potentials. The physical motivation for this was touched upon in 1.1. Rewards for separating them in the mathematical treatment are the possibility of incorporating confining potentials directly into the density-functional framework, as well as an improved understanding of the stability of Lieb’s $L^\infty + L^{3/2}$ potentials, and continuity of $E(v)$. We will see that Eq. (1.3) is not true, since the right-hand side is necessarily convex, but F°

is not. The defect is easily corrected by replacing F° by the mixed-state internal energy F . Furthermore, the maximization need be carried out only over potentials which are linear combinations of indicator functions of cells in the partition hierarchy \mathfrak{P} . This is a new result. Appendix A gives a refresher on basic functional analysis. In these notes, we freely use simple methods of infinitesimal (or nonstandard) analysis. Infinitesimal methods are a way of doing mathematics, and especially of dealing with mathematical idealizations, that ought to be familiar to physical scientists. Appendix B gives an account of what we need. It is relegated to an appendix simply because it is not the real subject of these notes, but only a tool. The reader may want to take an early look, since it begins to be used already in 2.3.1.

1.3. Particle types. We deal throughout with a system of \mathcal{N} indistinguishable particles. A condensed matter physicist or chemist generally thinks about DFT only as a theory of electrons. Our default assumption is indeed that the particles are fermions, but in principle, the theory could be applied to many different kinds of particles. Statistics only appear in the explicit construction of a wavefunction in 4.3.2, so for bosons a better bound on F° would be achievable. Spin is nothing but an annoyance here; it is always summed over and makes no qualitative difference. Finally, although the mutual interaction of most interest is the Coulomb interaction, the crucial ingredient is relative form-boundedness by kinetic energy. We do use positivity of the interaction in the form $F \geq 0$, but that is a matter of convenience and could be relaxed.

2. INTERNAL ENERGY AS FUNCTION OF STATE (\mathcal{E})

2.1. Kinetic energy. Throughout, the Hilbert space of \mathcal{N} -particle wavefunctions of appropriate symmetry is denoted by \mathcal{H} . In units such that $\hbar^2/2m$ has value 1, the kinetic energy of the wavefunction $\psi(\underline{x}, \underline{\sigma})$ is

$$\mathcal{E}_0(\psi) := \langle \psi | T | \psi \rangle = \int |\nabla \psi|^2 = \|\nabla \psi\|_2^2 \quad (2.1)$$

Remark 2.1 (integration and configuration variable conventions). \underline{x} denotes the full set of \mathcal{N} -particle configuration space coordinates $(x_1, x_2, \dots, x_{\mathcal{N}})$, and similarly $\underline{\sigma}$, the collection of spin coordinates. The integral in (2.1) illustrates some conventions that we will use. Integrals are over all available variables unless a restriction is indicated by a subscript on the integral sign. For example, $\int_{x_1=x} \dots$ means that the position coordinate of particle 1 is held fixed at x . The gradient symbol is subject to a *different* convention. The gradient in (2.1) is $3\mathcal{N}$ -dimensional; a subscript would indicate that differentiation is with respect to *only* the indicated coordinates. Integration over spin variables means summation.

In case ψ is in the domain of the operator T , so that $T\psi \in L^2(\mathbb{R})^3$ the quadratic form (2.1) is equal to the inner product of ψ with $T\psi$, which would be written $\langle \psi | T\psi \rangle$. The quadratic form, understood as a function $L^2(\mathbb{R}^3) \rightarrow \overline{\mathbb{R}}$ into the topped reals (see 2.3.1) $\overline{\mathbb{R}} := \mathbb{R} \cup \{+\infty\}$ has the advantage that it is meaningful for every ψ , and there are many wavefunctions satisfying $\mathcal{E}_0(\psi) < +\infty$ for which $T\psi$ does not exist. In general, Hamiltonians appear in DFT in the guise of expectation values. Quadratic form definitions are therefore appropriate.

2.2. Interaction energy. Now we add the interaction energy of the particles.

$$\mathcal{E}_\lambda(\psi) = \mathcal{E}_0(\psi) + \langle \psi | W | \psi \rangle. \quad (2.2)$$

Usually, the interaction W is a Coulomb interaction, $\sum_{i < j} |x_i - x_j|^{-1}$, but we consider generally what properties we want to require.

Call $a > 0$ an \mathcal{E}_0 form-bound for W if there is some $b \geq 0$ such that

$$\forall \psi \in L^2(\mathbb{R}^3), \quad |\langle \psi | W | \psi \rangle| \leq a\mathcal{E}_0(\psi) + b\|\psi\|^2. \quad (2.3)$$

Then, if a_0 is the infimum of all \mathcal{E}_0 form-bounds for W , says that W is \mathcal{E}_0 -form-bounded with relative bound a_0 . The relative form bound of the Coulomb interaction is zero, as shown in 2.2.1. For any $a > 0$, there is some constant b such that W_{Coul} is dominated by $\mathcal{E}_0 + b$. This fact is crucial to the stability of non-relativistic one-electron atoms of arbitrary Z . The motif of dominating an attractive

energy with a repulsive one and its significance for stability will recur. In the following, we assume that the interaction W is non-negative and \mathcal{E}_0 bounded with relative form bound zero (“Kato tiny relative to the kinetic energy”). Then, for some constants $c, d > 0$,

$$d\|\nabla\psi\|_2 \leq E(\psi) \leq c\|\psi\|_{H^1}. \quad (2.4)$$

2.2.1. Hardy inequality. The proof that the Coulomb interaction is Kato tiny relative to kinetic energy uses the following inequality.

Lemma 2.1 (Hardy’s inequality). *For $\psi \in H^1(\mathbb{R}^3)$,*

$$\int_{\mathbb{R}^3} \frac{|\psi|^2}{4r^2} dx \leq \int_{\mathbb{R}^3} |\nabla\psi|^2 dx$$

Proof. (sketch) We demonstrate the inequality for *real*-valued $\psi \in C_c^\infty(\mathbb{R}^3)$. Real and imaginary parts separate, and an approximation argument extends it to $H^1(\mathbb{R}^3)$.

Define $\phi = r^{1/2}\psi$, and note that $\phi(0) = 0$. Now,

$$\begin{aligned} (\nabla\psi)^2 &= \left(-\frac{\nabla r}{r^{3/2}}\phi + r^{-1/2}\nabla\phi \right)^2 \geq \frac{1}{4r^3}\phi^2 - 2\left(\frac{\phi}{2r^{3/2}}r^{-1/2} \right) \left(\frac{\partial\phi}{\partial r} \right) \\ &= \frac{1}{4r^2}\psi^2 - \frac{1}{2r^2} \frac{\partial\phi^2}{\partial r}, \end{aligned}$$

where in the inequality, we’ve thrown out a positive term. The second term in the final expression integrates to zero, since $\phi(x) = 0$ at $x = 0$ and for large x . \square

A couple of tricks are needed to apply the Hardy inequality to our situation. First, note that $1/r$ differs from $f(r|\epsilon) = (\text{if } r < \epsilon \text{ then } 1/r \text{ else } 0)$ by a bounded function and $f(r|\epsilon) \leq \epsilon/r^2$. Second, the expectation of $|x_i - x_j|^2$ can be handled by change of variables. Write the required integral as $\int |\psi|^2/|x_1 - x_2|^2 d(x_1 - x_2) dx_2 dx_3 \cdots dx_N$, and notice that $|\nabla_{x_1-x_2}\psi|^2 \leq |\nabla\psi|^2$.

2.3. $\mathcal{E} : \mathcal{H} \rightarrow \overline{\mathbb{R}}$ is lower semicontinuous.

2.3.1. Lower semicontinuity and extended-real-valued functions. We will often be interested in functions taking values in the topped reals

$$\overline{\mathbb{R}} := \mathbb{R} \cup \{+\infty\}. \quad (2.5)$$

A neighborhood of $+\infty$ contains an interval of the form $(a, \infty]$ for some $a \in \mathbb{R}$. Infinitesimally, this means that “ $x \simeq +\infty$ ” is to be understood literally. ${}^*\overline{\mathbb{R}}$ consists of ${}^*\mathbb{R}$ together with the new point $+\infty$. If $x \in {}^*\mathbb{R}$ is illimited, then ${}^\circ x = +\infty$.

Continuity of a function $f : X \rightarrow \mathbb{R}$ at a point $x \in X$ is defined this way: Given a tolerance ϵ , there is a neighborhood U of x such that $|f(y) - f(x)| < \epsilon$ for every $y \in U$. If we split the condition into $f(y) > f(x) - \epsilon$ and $f(y) < f(x) + \epsilon$, then the first characterizes lower semicontinuity at x , and the latter, upper semicontinuity. In case f is a map into $\overline{\mathbb{R}}$, we extend this definition for a point x with $f(x) = +\infty$. f is lower semicontinuous at x if for any $M \in \mathbb{R}$, there is a neighborhood U of x such that $f(y) > M$ whenever $y \in U$. In an infinitesimal idiom, $f : X \rightarrow \mathbb{R}$ is lower semicontinuous (abbreviated ‘lsc’) at x if $f(y) \gtrsim f(x)$ for all y in the halo of x . Equivalently, ${}^\circ f(y) \geq f(x)$. If f is lsc everywhere, then we call it simply “lower semicontinuous”. This is the property we are usually interested in.

Definition 2.1. A \mathbb{R} -valued function f on a topological space X is lower semicontinuous if $f^{-1}(-\infty, a]$ is closed for every $a \in \mathbb{R}$.

We usually work with the infinitesimal characterization, but there is also a useful geometrical intuition attached to the stated definition.

Definition 2.2. If $f : X \rightarrow \overline{\mathbb{R}}$, then its effective domain is

$$\text{dom } f := \{x \in X : f(x) < \infty\}. \quad (2.6)$$

The epigraph of f is the set on or above its graph in $X \times \mathbb{R}$:

$$\text{epi } f := \{(x, a) \in X \times \mathbb{R} : f(x) \leq a\}. \quad (2.7)$$

This implies that if $f(x) = \infty$, then there are no points in $\text{epi } f$ of the form (x, a) . Implicitly, $X \times \mathbb{R}$ is equipped with the product topology, so that $(y, b) \simeq (x, a)$ if and only if both $y \simeq x$ and $b \simeq a$.

We have introduced the notion of epigraph here because it provides another way to look at lower semicontinuity. $f : X \rightarrow \overline{\mathbb{R}}$ is lower semicontinuous if and only if its epigraph is closed. This perspective also gives a nice proof of the fact that the pointwise supremum of any set of lower semicontinuous functions is also lsc. Let $g = \sup_{\alpha \in I} f_\alpha$, with all the f_α 's lsc. Then, $\text{epi } g = \cap_{\alpha} \text{epi } f_\alpha$. But, the intersection of any collection of closed sets is closed, hence g is lsc.

2.3.2. Lsc theorem.

Theorem 2.1. $\mathcal{E} : \mathcal{H} \rightarrow \overline{\mathbb{R}}$ is lower semicontinuous.

Proof. Since $\mathcal{E} \sim \mathcal{E}_0$, it suffices to prove this for \mathcal{E}_0 . And that is very easily done by Fourier transformation:

$$\mathcal{E}_0(\psi) = \|\nabla \psi\|_2^2 = \int q^2 |\hat{\psi}(q)|^2 dq = \lim_{Q \rightarrow \infty} p_Q(\psi),$$

where

$$p_Q(\psi) := \int_{|q| \leq Q} q^2 |\hat{\psi}(q)|^2 dq.$$

But, $p_Q : \mathcal{H} \rightarrow \mathbb{R}$ is clearly continuous, and $p_Q(\psi)$ is monotonically non-decreasing in Q . As supremum of continuous functions, then, \mathcal{E}_0 is lsc. \square

2.4. Extension to mixed states.

2.4.1. Standard and nearstandard mixed states. To the vector $\psi \in \mathcal{H}$ is associated the pure state $|\psi\rangle\langle\psi| \in \mathbf{State}^\circ$, which is insensitive to the phase of ψ . A mixed state is a probabilistic mixture of pure states, and has the normal form

$$\gamma = \sum_{i \in \mathcal{I}} |\psi_i\rangle\langle\psi_i|, \quad j \neq k \Rightarrow \langle\psi_j|\psi_k\rangle = 0, \quad \|\psi_1\| \geq \|\psi_2\| \geq \dots > 0, \quad \sum_{i \in \mathcal{I}} \|\psi_i\|^2 < \infty. \quad (2.8)$$

The index set \mathcal{I} here is either \mathbb{N} , or $[1, \dots, n]$ for some $n \in \mathbb{N}$. Note, the normal form is not uniquely specified in case of degenerate eigenvalues. On a few occasions, we have use for an alternate normal form

$$\gamma = \sum_{i \in \mathcal{I}} c_i |\hat{\psi}_i\rangle\langle\hat{\psi}_i|. \quad (2.9)$$

The two are related by $c_i = \|\psi_i\|^2$, so that the $\hat{\psi}_i$'s are unit vectors.

We can also think of the set of mixed states \mathbf{State} as a convex cone in the Banach space of trace-class operators equipped with the trace norm,

$$\|A\|_{\text{Tr}} = \sum \langle\phi_i|A|\phi_i\rangle, \quad \{\phi_i\} \text{ an ONB.} \quad (2.10)$$

For a positive trace-class operator, the trace norm is just the trace, and for a mixed state this is the generalization of norm-squared for a state vector.

For $\psi \in {}^*\mathcal{H}$, ${}^\circ\psi$ is the standard vector infinitely close (“near”) to it, if such exists, else it is undefined. The map $\psi \mapsto |\psi\rangle\langle\psi|$ from \mathcal{H} to \mathbf{State}° ought to be continuous, so st $|\psi\rangle\langle\psi| = |{}^\circ\psi\rangle\langle{}^\circ\psi|$. Continuing in this vein, to deduce the normal form of st γ without explicitly working with the trace norm, note that both addition in \mathbf{State} and Tr should be continuous. Combined with the standard part operation in \mathbf{State}° , we conclude that

$$\text{st } \gamma = \quad \text{if } {}^\circ(\sum_{\mathcal{I}} \|\psi_i\|^2) = \sum_{{}^\circ\mathcal{I}} \|{}^\circ\psi_i\|^2 \text{ then } \sum_{i \in {}^\circ\mathcal{I}} |{}^\circ\psi_i\rangle\langle{}^\circ\psi_i| \text{ else undefined.} \quad (2.11)$$

Thus, γ is remote if any of the ψ_i 's is so, or if its norm is illimited, or if an appreciable weight is carried by infinitesimal ψ_i 's, or if the norm of γ is illimited (the if-clause fails because the first standard part is undefined).

Definition 2.3. The set of mixed states based on \mathcal{H} is denoted **State**, and equipped with the trace norm. The subset of pure states is denoted by \mathbf{State}° , and the normalized states by \mathbf{State}_1 . ‘State’ with no modifier means a member of **State**, and the distinction between a pure state and a corresponding state vector is often elided.

2.4.2. *Lower semicontinuity of \mathcal{E} .* With (2.11) and the nonstandard characterization of lower-semicontinuity, we can now extend Thm. 2.1 to **State**. However, it will not be used in the following.

Corollary 2.1. $\mathcal{E} : \mathbf{State} \rightarrow \overline{\mathbb{R}}$ is lower semicontinuous.

Proof. Take $\gamma \in \mathbf{State}$ a standard state and suppose $\gamma' = \sum_{\mathcal{I}} |\psi'_i\rangle\langle\psi'_i|$ is near γ . We need to show that $\mathcal{E}(\gamma') \gtrsim \mathcal{E}(\gamma)$. Since \mathcal{E} is non-negative and lsc on \mathbf{State}° ,

$${}^\circ\mathcal{E}(\gamma') \geq \text{st} \sum_{i=1}^n \mathcal{E}(\psi'_i) = \sum_{i=1}^n {}^\circ\mathcal{E}(\psi'_i) \geq \sum_{i=1}^m \mathcal{E}(\psi_i).$$

If \mathcal{I} is finite, just take $n = |\mathcal{I}|$, otherwise, the limit $n \rightarrow \infty$ yields ${}^\circ\mathcal{E}(\gamma') \geq \mathcal{E}(\gamma)$. \square

3. DENSITY (ρ)

The one-particle density associated with an \mathcal{N} -particle wavefunction ψ is

$$(\text{dens } \psi)(x) = \mathcal{N} \int_{x_1=x} |\psi|^2. \quad (3.1)$$

The right-hand side here is a linear function of the rank-one operator $|\psi\rangle\langle\psi|$ and thus has an immediate linear extension to mixed states:

$$\text{dens} \left(\sum c_i |\psi_i\rangle\langle\psi_i| \right) = \sum c_i \text{dens } \psi_i. \quad (3.2)$$

We will overload the notation ‘dens’ by writing $\text{dens } \psi$ for $\text{dens}(|\psi\rangle\langle\psi|)$.

3.1. **dens is continuous.** Integrating the previous display over x , one finds

$$\|\text{dens } \gamma\|_1 = \mathcal{N} \|\gamma\|_{tr}. \quad (3.3)$$

Densities are, of course, non-negative integrable functions. We use the notations

Definition 3.1.

$$\begin{aligned} \text{Dens} &= \{ \rho \in L^1(\mathbb{R}^3) : \rho \geq 0 \}, \\ \text{Dens}_{\mathcal{N}} &= \left\{ \rho \in \text{Dens} : \int \rho dx = \mathcal{N} \right\}. \end{aligned} \quad (3.4)$$

Theorem 3.1. $\text{dens} : \mathbf{State} \rightarrow \text{Dens}$ is continuous.

Proof. (3.3) shows that dens is actually a bounded linear map. \square

3.2. Tameness and nearstandardness.

3.2.1. *Tameness.* A density in $\text{Dens}_{\mathcal{N}}$ always integrates to \mathcal{N} . Thus, given a density ρ , there is always some function (many in fact) $R(\epsilon) : (0, N] \rightarrow \mathbb{R}^+$ such that $\int_{|x| \geq R(\epsilon)} \rho dx < \epsilon$ for all ϵ . When a set S of densities can all be controlled in this way by the same falloff function, S would be said to be *tight*. In a nonstandard context, we borrow this term to apply to a single \ast -density in $\ast\text{Dens}_{\mathcal{N}}$. By Leibniz’ Principle, $\rho \in \ast\text{Dens}_{\mathcal{N}}$ has a \ast -falloff function. We say it is tight if $R(\epsilon)$ is limited for non-infinitesimal ϵ . Here is an alternative phrasing.

Definition 3.2. A \ast -density ρ is *tight* if $\int_{|x| \geq R} \rho dx \simeq 0$ for every illimited R . A \ast -state γ is *tame* if it has limited internal energy and norm, and $\text{dens } \gamma$ is tight. A \ast -state which is not tame is *wild*.

3.2.2. *A fundamental spatial approximation.* Our approximation scheme has two ingredients, a spatial truncation method and a family of finite-dimensional subspaces of \mathcal{H} . Start with the former. Let $0 \leq \eta(x) \leq 1$ be a continuously differentiable cutoff function equal to one in the cubical box $\square(1/2) = [-1/2, 1/2]^3$ and supported in $\square(1) = [-1, 1]^3$, for example, a product of three 1D cutoff functions, and for $R > 0$, let $(\eta_R)(x) = \eta(x/R)$. Then, define the \mathcal{N} -particle cutoff operator by

$$(\Lambda_R \psi)(\underline{x}, \underline{\sigma}) = \left(\prod_{i=1}^{\mathcal{N}} \eta_R(x_i) \right) \psi(\underline{x}, \underline{\sigma}). \quad (3.5)$$

Thus, the state $\Lambda_R \psi$ is supported in the \mathcal{N} -particle box $\square^{\mathcal{N}}(R)$, which is to say that the density is zero outside $\square(R)$. $\Lambda_R \psi$ agrees with ψ inside $\square^{\mathcal{N}}(\frac{R}{2})$ and is an attenuated version outside, so we can get an upper bound on the norm of the difference by just integrating $|\psi|^2$ over the complement:

$$\|\Lambda_R \psi - \psi\|^2 \leq \int_{\square^{\mathcal{N}}(\frac{R}{2})^c} |\psi|^2 \leq \int_{\square(\frac{R}{2})^c} \text{dens } \psi \quad (3.6)$$

Bounding the energy of $\Lambda_R \psi$ is a little harder. Using $\nabla(\Lambda_R \psi) = \Lambda_R \nabla \psi + (\nabla \Lambda_R) \psi$ and $|a + b|^2 \leq 2|a|^2 + 2|b|^2$, we find

$$\begin{aligned} \mathcal{E}_0(\Lambda_R \psi) &\leq \mathcal{E}_0(\psi) + \int_{[\square^{\mathcal{N}}(\frac{R}{2})]^c} |\nabla \psi|^2 d\underline{x} + 2 \int_{[\square^{\mathcal{N}}(\frac{R}{2})]^c} |\nabla \Lambda_R|^2 |\psi|^2 d\underline{x} \\ &\leq 2\mathcal{E}_0(\psi) + 2 \int_{\square(\frac{R}{2})^c} |\nabla \eta_R|^2 (\text{dens } \psi) \\ &\leq 2\mathcal{E}_0(\psi) + \frac{\|\nabla \eta_1\|_{\infty}^2}{R^2} \int_{\square(\frac{R}{2})^c} \text{dens } \psi. \end{aligned} \quad (3.7)$$

To the box $\square(R)$ we associate two closed subspaces of \mathcal{H} . Let $\{\varphi_1, \varphi_2, \dots\}$ be the *single-particle* particle-in-a-box eigenstates in $\square(R)$. From this single-particle basis, construct a complete product basis (adding spin and antisymmetrizing as needed) for $\mathcal{H}(\square(R))$, the Hilbert space for \mathcal{N} particles in the box. \mathcal{H}_R is the subspace spanned by basis vectors with kinetic energy not exceeding R ; it is a *finite-dimensional* Hilbert space (asymptotically, $\dim \mathcal{H}_R$ does not grow any faster than $R^{9\mathcal{N}/2}$). Orthogonal projection $\mathcal{H}(\square(R)) \rightarrow \mathcal{H}_R$ is denoted by π_R . Now, if ψ has density supported in $\square(R)$, then $\psi \in \mathcal{H}(\square(R))$, and if, in addition, $\mathcal{E}_0(\psi) < \epsilon R$, then $\|\psi - \pi_R \psi\|^2 < \epsilon$. We use the notation Reg_R to denote the composition of smooth truncation followed by projection:

$$\text{Reg}_R := \pi_R \circ \Lambda_R. \quad (3.8)$$

Now, if ψ is tight and R is illimited, $\Lambda_R \psi \simeq \psi$ is immediate. The inequality (3.7) shows that also $\mathcal{E}_0(\Lambda_R \psi) \lesssim \mathcal{E}_0(\psi)$.

3.2.3. *Tame implies nearstandard.* Now we put the constructions of 3.2.2 to work. Two observations are critical. The constructions were described for $R \in \mathbb{R}_{>0}$, but work just as well for $R \in {}^*\mathbb{R}$. Also, since \mathcal{H}_R is finite-dimensional, with some finite orthonormal basis $\{\varphi_1, \dots, \varphi_n\}$, the same is true of ${}^*\mathcal{H}_R$. The only difference is that $\phi \in {}^*\mathcal{H}_R$ has hypercomplex coefficients and may have illimited norm. But, if $\phi = \sum \alpha_i {}^*\varphi_i$ is limited, it is nearstandard with standard part $\sum {}^\circ \alpha_i \varphi_i$.

The translation into standard terms of an external assertion to the effect that a property implies nearstandardness is a claim about relative compactness. For instance, the following proposition is essentially a variation of the Rellich-Kondrashov theorem. We will not, however, explicitly use the notion of compactness in these notes.

Proposition 3.1. *If $\psi \in {}^*\text{State}^\circ$ is tame, then ψ is nearstandard. Furthermore, in that case, $\text{dens st } \psi = \text{st dens } \psi$ and $\mathcal{E}(\text{st } \psi) \leq \text{st } \mathcal{E}(\psi)$.*

Proof. \mathcal{H} is a *complete* metric space, hence it suffices to show that ψ can be approximated to any standard accuracy by a (standard) vector in \mathcal{H} .

Since ψ is tight, $\Lambda_R \psi \simeq \psi$ for illimited R . Furthermore, according to inequality (3.7), $\Lambda_R \psi$ also has limited internal energy. Thus, the projection π_R will affect it only infinitesimally: $\text{Reg}_R \psi \simeq \psi$.

Now let $\epsilon > 0$ be given. The set

$$S(\epsilon) = \{a \in {}^*\mathbb{R} : \|\text{Reg}_a \psi - \psi\| < \epsilon\}$$

is internal, but it contains all illimited hyperreals. Thus, it must contain some limited R . $\text{Reg}_R \psi$ is a limited vector in ${}^*\mathcal{H}_R$, and therefore nearstandard by the remarks immediately preceding the proposition. Thus, ψ is within 2ϵ of a standard vector. But ϵ was arbitrary.

The last two statements of the Proposition now follow by Thms. 2.1 and 3.1. \square

Following the familiar pattern, we now extend Prop. 3.1 to mixed states. This is not trivial and will involve one of the trickiest pieces of reasoning in these notes.

Corollary 3.1. *If $\gamma \in \text{State}_1$ is tame, then γ is nearstandard. Consequently, $\text{dens st } \gamma = \text{st dens } \gamma$ and $\mathcal{E}(\text{st } \gamma) \leq {}^\circ \mathcal{E}(\gamma)$.*

Proof. Tackle this in contrapositive form. There are three ways γ can be remote. If $\|\gamma\|_{\text{Tr}}$ is illimited, γ is certainly wild, and if $\|\psi_i\|^2$ is appreciable for some remote ψ_i in the support of γ , then γ is wild because Prop. 3.1 says that ψ_i is. The third possibility is that γ has a tail $\gamma' = \sum_{i \geq N} c_i |\psi_i\rangle\langle\psi_i|$ with $c_i \simeq 0$ for $i \geq N$ and $\sum_{i \geq N} c_i = M \gg 0$. The following Lemma 3.1 shows that γ'/M is wild. Hence, so is γ . \square

Lemma 3.1. *Suppose the mixed state*

$$\gamma = \sum c_i |\hat{\psi}_i\rangle\langle\hat{\psi}_i| \in {}^*\text{State}_1$$

has all coefficients infinitesimal: $c_i \simeq 0$. Then, γ is wild.

Proof. Assume to the contrary that γ is tame, so that $\mathcal{E}(\gamma) = E \ll \infty$ and $\rho := \text{dens } \gamma$ is tight. Define the *weight* of an internal set $B \subset {}^*\mathbb{N}$ to be $\text{wt } B = \sum_{i \in B} c_i$. Since the c_i are infinitesimal, if $\text{wt } B \gg 0$, then $|B|$ is illimited. We are going to derive a contradiction as follows. For some limited r , we obtain an *illimited* set of indices $A(r)$ such that $\|\text{Reg}_r \hat{\psi}_i - \hat{\psi}_i\| < \epsilon$ for $i \in A(r)$. Then $\|\text{Reg}_r \hat{\psi}_i\| > 1 - \epsilon > 1/2$, and since the $\hat{\psi}_i$'s are orthonormal, a quick calculation reveals that $\langle \text{Reg}_r \hat{\psi}_i | \text{Reg}_r \hat{\psi}_j \rangle \leq (2 + \epsilon)/\epsilon < 3\epsilon$. Thus, the cosine of the angle between $\text{Reg}_r \hat{\psi}_i$ and $\text{Reg}_r \hat{\psi}_j$ is less than 12ϵ . For small ϵ , this condition is a restriction; a set of directions in the *limited-dimensional* space \mathcal{H}_r satisfying it must be limited, but $A(r)$ is not. Now we proceed to produce the required r and $A(r)$. Let standard $\epsilon \in (0, 1/12)$ be given, and define for $R \in {}^*\mathbb{R}$, the *internal* set

$$A(R) := \left\{ i : \int_{\square(R/2)^c} \text{dens } \hat{\psi}_i \leq \frac{1}{\epsilon} \int_{\square(R/2)^c} \rho \text{ and } \mathcal{E}_0(\hat{\psi}_i) \leq \frac{E}{\epsilon} \right\}.$$

For every R , $\text{wt } A(R) \geq 1 - \epsilon$, for if this failed, there would be too much mass outside $\square(R/2)$ or $\mathcal{E}(\gamma)$ would exceed E . Consequently, $|A(R)| > (1 - \epsilon)/c_1 \simeq \infty$, independently of R .

If R is illimited, then by (3.6, 3.7) and tightness of ρ , $\Lambda_R \hat{\psi}_i \simeq \hat{\psi}_i$ and $\mathcal{E}_0(\Lambda_R \hat{\psi}_i) \ll \infty$ for $i \in A(R)$. And, this implies $\|\text{Reg}_R \hat{\psi}_i - \hat{\psi}_i\| \simeq 0$. As a result, the *internal* set

$$C := \{R \in {}^*\mathbb{R} : \forall i \in A(R), \|\text{Reg}_R \hat{\psi}_i - \hat{\psi}_i\| < \epsilon\}$$

contains all illimited hyperreals. By underflow, C contains some $r \ll \infty$. $A(r)$ is the hyperfinite set discussed in the first paragraph. \square

4. INTERNAL ENERGY AS FUNCTION OF DENSITY (F° AND F)

4.1. Constrained energy minimization. We now move away from talking directly, or at least primarily, about states, instead addressing them through the mediation of density. The discussion of the Levy-Lieb constrained search formulation in 1.2 introduced the idea of minimizing internal energy over all states with given density. However, it is now clear that we want to consider two

different internal energy functions of density, depending upon whether all states or only pure states are considered. Thus, we define

$$\begin{aligned} F^\circ(\rho) &= \inf \{ \mathcal{E}(\psi) : \psi \in \text{State}^\circ, \text{dens } \psi = \rho \} \\ F(\rho) &= \inf \{ \mathcal{E}(\gamma) : \gamma \in \text{State}, \text{dens } \gamma = \rho \}. \end{aligned} \quad (4.1)$$

Important questions about these functions are these: For which densities they are finite? Are there useful upper/lower bounds in terms of explicit functions of density? Are F° and F lower semicontinuous? Are the infima in their definitions actually realized (are they actually minima)? Using the results of 2.3 and 2.4, we can immediately settle the last question.

Theorem 4.1. *The infima in the definitions of F° and F are realized.*

Proof. Let standard ρ be given and assume $F^\circ(\rho) < +\infty$, as otherwise there is nothing to prove. Let $\psi \in {}^*\mathcal{H}$ be such that $\mathcal{E}(\psi) \simeq {}^*F^\circ({}^*\rho)$. Since ψ is thus tame, it is nearstandard by Thm. 3.1 so $\text{st } \psi$ exists, and according to Thm. 2.1, $\mathcal{E}(\text{st } \psi) \lesssim \mathcal{E}(\psi) \simeq F^\circ(\rho)$. Since $F^\circ(\rho)$ and $\mathcal{E}(\text{st } \psi)$ are both standard, they are equal.

The proof for F is entirely analogous, just substituting Cor. 2.1 for Thm. 2.1, and Cor. 3.1 for Thm. 2.1. \square

4.2. Sobolev norms and kinetic energy bounds. In this section, we derive some important lower bounds on F° and F .

4.2.1. Basic inequality. Starting from (3.1), we deduce

$$\begin{aligned} \nabla \rho(x) &\leq 2\mathcal{N} \int_{x_1=x} |\psi^* \nabla_1 \psi|^2 \leq 2\mathcal{N} \left(\int_{x_1=x} |\psi|^2 \right)^{1/2} \left(\int_{x_1=x} |\nabla_1 \psi|^2 \right)^{1/2} \\ &= 2\rho(x)^{1/2} \left(\int_{x_1=x} |\nabla_1 \psi|^2 \right)^{1/2}. \end{aligned} \quad (4.2)$$

The subscript on ∇ indicates that differentiation is only with respect to x_1 .

4.2.2. Two bounds. There are two useful things we can do with this result. First, squaring, dividing through by ρ and noting that $|\nabla \rho|^2 / \rho = |\nabla \sqrt{\rho}|^2$,

$$\|\nabla \sqrt{\rho}\|_2^2 \leq \frac{4}{\mathcal{N}} \mathcal{E}_0(\psi). \quad (4.3)$$

To extend this to a mixed state $\gamma = \sum |\psi_i\rangle\langle\psi_i|$, abbreviate $\rho = \text{dens } \gamma$, $\rho_i = \text{dens } \psi_i$ and use $\nabla \rho = 2\sqrt{\rho} \nabla \sqrt{\rho}$ and the Cauchy-Schwartz inequality to obtain

$$|\sqrt{\rho} \nabla \sqrt{\rho}| = \left| \sum_i \sqrt{\rho_i} \nabla \sqrt{\rho_i} \right| \leq \left(\sum_i \rho_i \right)^{1/2} \left(\sum_i |\nabla \sqrt{\rho_i}|^2 \right)^{1/2}.$$

Since $\rho = \sum_i \rho_i$, divide through by $\sqrt{\rho}$ and square to get $|\nabla \sqrt{\rho}|^2 \leq \sum_i |\nabla \sqrt{\rho_i}|^2$. Integrating and inserting (4.3) yields

$$\|\nabla(\text{dens } \gamma)^{1/2}\|_2^2 \leq \frac{4}{\mathcal{N}} \mathcal{E}_0(\gamma). \quad (4.4)$$

For given ρ , this applies to any γ with $\text{dens } \gamma = \rho$ to give

$$\|\nabla \sqrt{\rho}\|_2^2 \leq \frac{4}{\mathcal{N}} F(\rho). \quad (4.5)$$

Since we wish to work with density rather than the square root of density, the inequality (4.4) is not always convenient. Alternatively, returning to (4.2), integrating and applying the Cauchy-Schwartz inequality,

$$\int |\nabla \rho| \leq 2 \left(\int \rho \right)^{1/2} \left(\frac{1}{\mathcal{N}} \int |\nabla \psi|^2 \right)^{1/2},$$

which yields $\|\nabla\rho\|_1 \leq 2\sqrt{\mathcal{E}_0(\psi)}$. This is extended to mixed states by a method similar to what gave (4.4), with the result

$$\|\nabla \text{dens } \gamma\|_1 \leq 2\sqrt{\mathcal{E}_0(\gamma)}. \quad (4.6)$$

Again, we turn this into a bound for F , as

$$\|\nabla\rho\|_1 \leq 2\sqrt{F(\rho)}. \quad (4.7)$$

4.3. Effective domain. In this section we will show that

$$\boxed{\text{dom } F^\circ = \text{dom } F = \mathcal{J}_\mathcal{N} := \{\rho \in \text{Dens}_\mathcal{N} : \sqrt{\rho} \in H^1(\mathbb{R})^3\}} \quad (4.8)$$

is the effective domain of both F° and F . That $\mathcal{J}_\mathcal{N} \supseteq \text{dom } F \supseteq \text{dom } F^\circ$ follows from the results of 4.2 and the fact that $F \leq F^\circ$. So what needs to be shown is $\mathcal{J}_\mathcal{N} \subset \text{dom } F^\circ$. This is done by an explicit construction.

4.3.1. Transporting densities and wavefunctions. Suppose a density ρ_0 is given on some domain Ω and we wish to pull back the mass in Ω along a bijection $\varphi : \mathbb{R}^3 \rightarrow \Omega$. That is, we want the density ρ on \mathbb{R}^3 such that

$$\int_A \rho = \int_{\varphi(A)} \rho_0$$

for every measurable $A \subset \mathbb{R}^3$. The solution to this problem is very familiar: with $J(x) = |\partial\varphi/\partial x|$ denoting the Jacobian determinant and abbreviating $y = \varphi(x)$,

$$\rho(x) = \rho_0(\varphi(x)) J(x) = \rho_0(y) J(x). \quad (4.9)$$

What if, instead, we are given a wavefunction ψ_0 on $\Omega^\mathcal{N}$ with $\text{dens } \psi_0 = \rho_0$ and wish to find ψ with $\text{dens } \psi = \rho$? A solution is nearly as easy. First extend the map φ to the \mathcal{N} -particle configuration space as

$$\underline{y} = \underline{\varphi}(\underline{x}) = (\varphi(x_1), \varphi(x_2), \dots, \varphi(x_\mathcal{N})).$$

Then, we can take

$$\psi(\underline{x}) = \psi_0(\underline{y}) \left| \frac{\partial \varphi}{\partial \underline{x}} \right|^{\frac{1}{2}} = \psi_0(\underline{y}) \prod_i J(x_i)^{1/2}. \quad (4.10)$$

Naturally, one wants to bound $\|\psi\|_{H^1(\mathbb{R}^3)}$ in terms of $\|\psi_0\|_{H^1(\Omega)}$, but we shall not pursue that problem in general, but move on to the special case that interests us.

4.3.2. Wavefunction construction for $\rho \in \mathcal{J}_\mathcal{N}$. We now construct [1, 9, 10] a wavefunction ψ with $\text{dens } \psi$ equal to given ρ in $\mathcal{J}_\mathcal{N}$, and calculate a bound on its internal energy. Using 4.3.1 as inspiration, we find a map $\varphi : \mathbb{R}^3 \rightarrow \Omega$ onto region Ω and a type of density ρ_0 on that region for which an appropriate Slater determinant can be easily found. We leave spin coordinates out of consideration, so implicitly all σ are supposed to take the maximum value.

Take $\Omega = [0, 1] \times \mathbb{R} \times \mathbb{R}$, and assume ρ_0 has the property

$$\int \rho_0(y^1, y^2, y^3) dy^2 dy^3 = \mathcal{N}. \quad (4.11)$$

Then, it is readily apparent that the one-particle wavefunctions $[y \equiv (y^1, y^2, y^3)]$

$$\phi_{0,k}(y) := \left(\frac{\rho_0(y)}{\mathcal{N}} \right)^{1/2} e^{2\pi i k y^1} \quad (4.12)$$

indexed by integer k are orthonormal over Ω . Since the density for each ϕ_k is $\rho_0(y)/\mathcal{N}$, a Slater determinant of ϕ_k 's for \mathcal{N} distinct k 's will give density ρ_0 .

All that is now required is a mapping of \mathbb{R}^3 onto Ω which carries the given ρ on to a ρ_0 satisfying (4.11). With the definition

$$J(x^1) := \frac{1}{\mathcal{N}} \int \rho(x) dx^2 dx^3, \quad (4.13)$$

one immediately checks that the following works.

$$\begin{aligned}\varphi(x) &= (y^1, y^2, y^3) = (y^1, x^2, x^3), \\ y^1 &= \int_{-\infty}^{x^1} J(z^1) dz^1.\end{aligned}\tag{4.14}$$

Mass is being redistributed only along the first direction, at a varying rate so that (4.11) is satisfied.

Of course, the notation was chosen because J actually is the Jacobian of φ :

$$J(x^1) = \frac{dy^1}{dx^1} = \left| \frac{\partial y}{\partial x} \right| = \frac{\rho(x)}{\rho_0(y)}.\tag{4.15}$$

According to (4.9), the $\phi_{0,k}$ of (4.12) are now transported to \mathbb{R}^3 as

$$\phi_k(x) = \phi_{0,k}(y) \sqrt{J(x)} = \sqrt{\frac{\rho(x)}{\mathcal{N}}} e^{2\pi i k y^1}.\tag{4.16}$$

Without further calculation, the general construction assures us that the ϕ_k are orthonormal on \mathbb{R}^3 and that a Slater determinant $\psi = |\phi_{k_1}, \dots, \phi_{k_{\mathcal{N}}}|$ of them yields the desired density ρ . As illustrated here, y may be regarded as a function of x or vice-versa, and either one as a coordinatization of \mathbb{R}^3 or of Ω as convenient, because the correspondence between x and y is bijective.

The wavefunction now being constructed, all that remains is to bound on the kinetic energy of ψ by bounding the kinetic energy of each ϕ_k :

$$\mathcal{E}_0(\phi_k) = \frac{1}{\mathcal{N}} \|\nabla \sqrt{\rho}\|_2^2 + \frac{(2\pi k)^2}{\mathcal{N}} \int \rho(x) \left(\frac{dy^1}{dx^1} \right)^2 dx.\tag{4.17}$$

The second term here will occupy us with a bit of Cauchy-Schwartz calisthenics. Using (4.15),

$$\int \left(\frac{dy^1}{dx^1} \right)^2 \rho(x) dx = \int \rho_0(y) J^2 dy = \mathcal{N} \int_0^1 J^2 dy^1.\tag{4.18}$$

Estimate the integrand in the final expression as

$$\begin{aligned}J(x^1)^2 &= \left(\int_{-\infty}^{x^1} J(z)^{\frac{1}{2}} \left(J(z)^{-\frac{1}{2}} \frac{dJ(z)}{dz} \right) dz \right)^2 \leq \left(\int J dx^1 \right) \int J^{-1} \left| \frac{dJ}{dx^1} \right|^2 dx^1 \\ &= \int J^{-1} \left| \frac{dJ}{dx^1} \right|^2 dx^1.\end{aligned}\tag{4.19}$$

The first equality is just the fundamental theorem of calculus, the inequality is Cauchy-Schwartz, and the final equality follows from $\int J dx^1 = \int dy^1 = 1$. To evaluate the final integral, Again, we go to work on the integrand with the Cauchy-Schwartz inequality, to obtain

$$\left| \frac{dJ}{dx^1} \right|^2 = \left(\frac{1}{\mathcal{N}} \int 2\sqrt{\rho(x)} \left| \frac{\partial \sqrt{\rho(x)}}{\partial x^1} \right| dx^2 dx^3 \right)^2 \leq \frac{4}{\mathcal{N}^2} J(x^1) \int |\nabla \sqrt{\rho}|^2 dx^2 dx^3.$$

Inserting this into (4.19),

$$J(x^1)^2 \leq \int J^{-1} \left| \frac{dJ}{dx^1} \right|^2 dx^1 \leq \frac{4}{\mathcal{N}^2} \|\nabla \sqrt{\rho}\|_2^2$$

Finally, returning to (4.17) via (4.18) results in

$$\mathcal{E}_0(\phi_k) \leq \frac{1}{\mathcal{N}} \left(1 + \left(\frac{4\pi k}{\mathcal{N}} \right)^2 \right) \|\nabla \sqrt{\rho}\|_2^2.$$

This is for one of the orbitals entering our Slater determinant, so summing over \mathcal{N} values of $|k| < \mathcal{N}$,

$$F^\circ(\rho) \leq \mathcal{E}_0(\psi) \leq (1 + 16\pi^2) \|\nabla \sqrt{\rho}\|_2^2.\tag{4.20}$$

4.4. $\mathcal{J}_{\mathcal{N}} \subset L^1 \cap L^3 \cap W^{1,1}$. Combining the upper bound on $F^\circ(\rho)$ from (4.20) with the lower bound of (4.5) yields

$$\boxed{c\|\sqrt{\rho}\|_{H^1}^2 \leq F(\rho) \leq F^\circ(\rho) \leq c'\|\sqrt{\rho}\|_{H^1}^2} \quad (4.21)$$

for some \mathcal{N} -dependent constants c and c' . The demonstration of (4.8) is thus achieved, with some extra information.

Characterizations in terms of norms of ρ rather than of $\sqrt{\rho}$ can also be useful. Inequality (4.7) shows that $\mathcal{J}_{\mathcal{N}} \subset W^{1,1}$, where $W^{1,1}$ is the Sobolev space with norm $\|f\|_{W^{1,1}} = \int (|f| + |\nabla f|) dx$. Also[1], the Sobolev inequality

$$3\left(\frac{\pi}{2}\right)^{4/3} \|f\|_6^2 \leq \|\nabla f\|_2^2,$$

with $\sqrt{\rho}$ substituted for f yields

$$\|\rho\|_3 \leq cF(\rho), \quad (4.22)$$

for some constant c . Thus, $\mathcal{J}_{\mathcal{N}} \subset L^3$, as well. The inequality (4.22) will be applied to the consideration of potentials in §6.4.

4.5. Internal energy is lower semicontinuous on $\text{Dens}_{\mathcal{N}}$.

4.5.1. *Topologies on $\text{Dens}_{\mathcal{N}}$.* The subject of this section, lower semicontinuity of F° and F on $\text{Dens}_{\mathcal{N}}$, is important to the Fenchel conjugacy with the ground-state energy E discussed in §6, among other things. But, lower semicontinuous with respect to what topology on $\text{Dens}_{\mathcal{N}}$? In §3.1, we showed that $\text{dens} : \text{State} \rightarrow \text{Dens}$ is continuous with respect to the L^1 topology on Dens . We have not made use of that result, yet. Continuity of dens thus holds for any topology weaker than L^1 . That is, so long as the topology is not so weak that it fails to be Hausdorff, $\text{dens st } \gamma = \text{st dens } \gamma$ for nearstandard γ . (Recall that a Hausdorff topology is such that for $x \neq y$, one can find neighborhoods U of x and V of y , such that $U \cap V = \emptyset$.) The topology is implicit in the standard part operation, which is not well-defined for a non-Hausdorff topology. Continuity of dens is important for the proof that F° is lsc, but there is another condition that we need to impose.

Definition 4.1. A topology τ on $\text{Dens}_{\mathcal{N}}$ is *weak-but-not-leaky* if it is weaker than, or equivalent to, L^1 , yet Hausdorff and such that all nearstandard points are tight.

An obvious example of such a topology is the weak- L^1 topology induced by the seminorms $\rho \mapsto \int f\rho$ for $f \in L^\infty$. A disadvantage of the weak- L^1 topology is that it is not metrizable. In §5, we will construct a metrizable weak-but-not-leaky topology on $\text{Dens}_{\mathcal{N}}$, the weak- \mathfrak{P} topology, which is even weaker than weak- L^1 .

4.5.2. Pure-state internal energy.

Theorem 4.2. $F^\circ : (\text{Dens}_{\mathcal{N}}, \tau) \rightarrow \overline{\mathbb{R}}$ is lsc, if τ is weak-but-not-leaky.

Proof. Let $\rho \in \text{Dens}_{\mathcal{N}}$ be given and $\rho' \in {}^*\text{Dens}_{\mathcal{N}}$ be τ -near ρ . Assume $F^\circ(\rho') \ll \infty$, since otherwise there is nothing to show. ρ' is tight by definition of weak-but-not-leaky, so there is a tame ψ' with $\text{dens } \psi' = \rho'$, and $\mathcal{E}(\psi') \simeq F^\circ(\rho')$. By Prop. 3.1 and 4.5.1, ψ' is nearstandard, with $\mathcal{E}(\text{st } \psi') \lesssim \mathcal{E}(\psi') \simeq F^\circ(\rho')$, and $\text{dens st } \psi' = \text{st dens } \psi' = \text{st } \rho' = \rho$. \square

4.5.3. Mixed-state internal energy.

Definition 4.2. If A is a subset of a vector space, then its convex hull is defined as

$$\text{co } A := \left\{ \sum_{i=1}^n \alpha_i x_i : \sum_{i=1}^n \alpha_i = 1 \text{ and } x_1, \dots, x_n \in A \right\}. \quad (4.23)$$

The closed convex hull of A , denoted $\overline{\text{co}} A$ is the closure of $\text{co } A$ in whatever topology is being considered on the ambient vector space.

The goal now is to extend the result of 4.5.2 from F° to F by showing that $\text{epi } F = \overline{\text{co}} \text{epi } F^\circ$. For, lower semicontinuity (convexity) of F is equivalent to closure (resp., convexity) of its epigraph.

Theorem 4.3. $\text{epi } F = \overline{\text{co}} \text{epi } F^\circ$ in $\text{Dens}_{\mathcal{N}} \times \mathbb{R}$, with any weak-but-not-leaky topology on $\text{Dens}_{\mathcal{N}}$.

Proof. To show that $\text{epi } F = \overline{\text{co}} \text{epi } F^\circ$, we establish inclusion in both directions.

(a) $\text{epi } F \subseteq \overline{\text{co}} \text{epi } F^\circ$: It suffices to show that, if $\text{dens } \gamma = \rho$, then $(\rho, \mathcal{E}(\gamma)) \in \overline{\text{co}} \text{epi } F^\circ$. From the second normal form for γ , define

$$w_n = \sum_{i=1}^n c_i, \quad \gamma_n = \sum_{i=1}^n \frac{c_i}{w_n} |\hat{\psi}_i\rangle\langle\hat{\psi}_i|.$$

Then,

$$\mathcal{E}(\gamma_n) = \sum_{i=1}^n \frac{c_i}{w_n} \mathcal{E}(\hat{\psi}_i) \rightarrow \mathcal{E}(\gamma), \quad \text{dens } \gamma_n = \sum_{i=1}^n \frac{c_i}{w_n} \text{dens } \hat{\psi}_i \rightarrow \text{dens } \gamma.$$

$(\text{dens } \gamma_n, \mathcal{E}(\gamma_n))$ is in $\text{coepi } F^\circ$, because $(\text{dens } \hat{\psi}_i, \mathcal{E}(\hat{\psi}_i)) \in \text{epi } F^\circ$.

Therefore, $(\text{dens } \gamma, \mathcal{E}(\gamma)) \in \overline{\text{co}} \text{epi } F^\circ$, as required.

(b) $\text{epi } F \supseteq \overline{\text{co}} \text{epi } F^\circ$: Let $(\rho, \lambda) \in \overline{\text{co}} \text{epi } F^\circ$ be given (standard). If $\lambda = +\infty$, there is nothing to show, so assume $\lambda < +\infty$. By definition of closed convex envelope, there is $N \in {}^*\mathbb{N}$ and collections of coefficients and orthonormal $*$ -pure-states such that

$$\sum_{i \leq N} c_i = 1, \quad \rho \simeq \sum_{i \leq N} c_i \text{dens}(\psi_i), \quad \lambda \simeq \sum_{i \leq N} c_i \mathcal{E}(\psi_i). \quad (4.24)$$

But this shows that $\gamma = \sum_{i=1}^N c_i |\psi_i\rangle\langle\psi_i| \in {}^*\text{State}_1$ satisfies $\text{dens } \gamma \simeq \rho$ and $\mathcal{E}(\gamma) \simeq \lambda$. Since γ is thus tame, Cor. 3.1 implies that $\mathcal{E}(\text{st } \gamma) \leq \lambda$ and $(\text{dens st } \gamma, \mathcal{E}(\text{st } \gamma)) \in \text{epi } F$. \square

This of course immediately gives us lower-semicontinuity.

Corollary 4.1. $F : (\text{Dens}_{\mathcal{N}}, \tau) \rightarrow \overline{\mathbb{R}}$ is lower semicontinuous if τ is weak-but-not-leaky.

5. WEAK- \mathfrak{P} TOPOLOGY ON $\text{Dens}_{\mathcal{N}}$

The aim in this section is to construct a convenient, metrizable, weak-but-not-leaky topology on $\text{Dens}_{\mathcal{N}}$ which is called the weak- \mathfrak{P} topology. Lower semicontinuity of F with respect to it immediately implies the same for any $\|\cdot\|_p$, $1 \leq p < \infty$. Although we are interested in only $\text{Dens}_{\mathcal{N}}$, the weak- \mathfrak{P} metric will be defined on all of Dens .

5.1. Hierarchy of partitions. Let \mathfrak{P}^0 denote a partition of \mathbb{R}^3 into a regular grid of cubes of side length 1, with corners at integer coordinates. To make it a genuine partition, we agree that each cell includes its front, left and bottom faces, but not the back, right or top. Now, divide each cell of \mathfrak{P}^0 into 2^3 cubes of side length 2^{-1} ; these latter cubes are the cells of \mathfrak{P}^1 . Continue subdividing so that each cell of \mathfrak{P}^n is the union of 2^3 cells of \mathfrak{P}^{n+1} to obtain an infinite hierarchy of partitions

$$\mathfrak{P}^0 < \mathfrak{P}^1 < \dots$$

A set A is \mathfrak{P}^n -measurable if it is the union of cells of \mathfrak{P}^n , \mathfrak{P}^n -finitely-measurable if it is a finite such union, and simply \mathfrak{P} -measurable if a finite union of cells. In this last case, A is actually \mathfrak{P}^m -finitely-measurable, where m is the maximal rank of members of A , but the terminology allows us to avoid mentioning m .

We define a linear projection π_n from integrable functions to \mathfrak{P}^n -measurable integrable functions as follows. For a function $f \in L^1(\mathbb{R}^3)$, $\pi_n f$ is constant over each cell of \mathfrak{P}^n , with value on a cell equal to the average of f over that cell. Thus, f and $\pi_n f$ have the same integral over each cell.

5.1.1. a metric for the weak- \mathfrak{P} topology.

Definition 5.1. For A a \mathfrak{P}^n -finitely-measurable set, the continuous seminorm p_A on $L^1(\mathbb{R}^3)$ is defined as

$$p_A(f) := \int_A |\pi_n f| \leq \|f\|_1. \quad (5.1)$$

The weak- \mathfrak{P} topology is simply the topology generated by these seminorms. However, not all these seminorms are needed. The smaller collection $\{p_\Omega : \Omega \in \mathfrak{P}\}$ suffices. That is, $f_n \xrightarrow{\text{wk-}\mathfrak{P}} f$ if and only if $p_\Omega(f - f_n) \rightarrow 0$ for every $\Omega \in \mathfrak{P}$. We can go even further and use just $\{p_{C_n}, n = 0, 1, \dots\}$ where C_n is the largest \mathfrak{P}^n -measurable set inside the ball $B_0(n)$ of radius n centered at 0. Finally, if a metric is preferred,

$$d_{\mathfrak{P}}(f, g) = \sum_{n=0}^{\infty} \frac{p_{C_n}(f - g)}{2^{n+2}} \quad (5.2)$$

generates the weak- \mathfrak{P} topology. Note that for $f, g \in L^1$, $d_{\mathfrak{P}}(f, g) \leq \|f - g\|_1/2$.

5.2. How weak is it? An example shows an important way in which weak- \mathfrak{P} is weaker than the L^1 topology. Let f equal 1 on some cell $\Omega \in \mathfrak{P}^0$, and 0 elsewhere. One can easily construct a function f_n which is zero off Ω , and inside Ω equal to 2 on half the \mathfrak{P}^n cells and to 0 on the other half in such a way that $\pi_m f_n \equiv 0$ for $m < n$. Then, $f_n \xrightarrow{\text{wk-}\mathfrak{P}} f$, but $\|f_n - f\|_p = 1$ for all n and all $1 \leq p < \infty$. On the other hand, since $|\int_\Omega f| \leq \|f\|_p |\Omega|^{1/p'}$, L^p convergence implies wk- \mathfrak{P} convergence. Thus, on $\text{Dens}_{\mathcal{N}}$, weak- \mathfrak{P} is weaker than L^p ($1 \leq p < \infty$).

The f_n of the example became more and more oscillatory as n increased. In our application, that sort of behavior implies ever increasing kinetic energy.

5.3. Approximation by projection.

5.3.1. Weak- \mathfrak{P}^n is weak-but-not-leaky. Recall that C_n is the union of \mathfrak{P}^n -cells contained fully within the ball of center 0 and radius n , and 1_{C_n} is the indicator function of C_n , equal to 1 on C_n and to zero on its complement C_n^c .

Lemma 5.1. *Let $g \in L^1(\mathbb{R}^3)$. Then, both $1_{C_n} \Pi_n g$ and $\Pi_n g$ tend to g in L^1 as $n \rightarrow \infty$.*

Proof. Consider first a continuous function with bounded support, $f \in C_c(\mathbb{R}^3)$. It is easy to see that $\|f - \Pi_n f\|_1 \rightarrow 0$. For, being compactly supported, f is uniformly continuous, so given $\epsilon > 0$, there is $\delta > 0$ such that $|f(x+y) - f(x)| < \epsilon/M$ whenever $|x-y| < \delta$, where M is the Lebesgue measure of the support of f . But then the average of f over a region A with diameter less than δ differs by less than ϵ/M from its value at any point in A . Therefore, if n is large enough that $\sqrt{3}2^{-n} < \delta$, it follows that $\|f - \Pi_n f\|_1 < \epsilon$.

Now, given $g \in L^1(\mathbb{R}^3)$, and $\epsilon > 0$, There is $f \in C_c(\mathbb{R}^3)$ with $\|g - f\|_1 < \epsilon/3$, because continuous functions with bounded support are dense in $L^1(\mathbb{R})$. Then, by the triangle inequality,

$$\begin{aligned} \|g - 1_{C_n} \Pi_n g\|_1 &\leq \|\rho - f\|_1 + \|f - \Pi_n f\|_1 + \|\Pi_n(f - \rho)\|_1 + \|\Pi_n 1_{C_n^c} g\|_1 \\ &\leq 2\|\rho - f\|_1 + \|f - \Pi_n f\|_1 + \|1_{C_n^c} g\|_1. \end{aligned}$$

The final line follows because Π_n is an L^1 contraction. In the final line, the second term tends to zero by the previous paragraph and the last term simply because g is integrable. Therefore, $\limsup_{n \rightarrow \infty} \|g - 1_{C_n} \Pi_n g\|_1 < \epsilon$. The proof for $\Pi_n g$ is even easier, as the final term in the displayed inequality is absent in that case. \square

This lemma immediately implies that the weak- \mathfrak{P} topology is weak-but-not-leaky, as shown in the following corollary. However, it may be skipped without loss since a stronger result is proven independently in Thm. 5.1.

Corollary 5.1. *The weak- \mathfrak{P} topology on $\text{Dens}_{\mathcal{N}}$ is weak-but-not-leaky.*

Proof. That the weak- \mathfrak{P} topology is weaker than L^1 is trivial. Lemma 5.1 shows that it is Hausdorff. There only remains to show that if $\rho \in \text{Dens}_{\mathcal{N}}$ and ρ' is $d_{\mathfrak{P}}$ -near ρ , that ρ' is tight. But, $\int_{C_n} |\Pi_n(\rho' - \rho)| \simeq 0$ implies that $\int_{C_n} \rho' \simeq \int_{C_n} \rho$. Since both integrate to \mathcal{N} , $\int_{C_n^c} \rho' \simeq \int_{C_n^c} \rho$ for all limited n . But $\int_{C_n^c} \rho'$ is decreasing in n and $\int_{C_n} \rho \rightarrow 0$. \square

5.4. Weak- \mathfrak{P} with internal energy control is as good as L^1 . The next theorem shows that if a sequence ρ_n converges with respect to $d_{\mathfrak{P}}$, then the only way it can avoid converging with respect to L^1 is to have divergent internal energy.

5.4.1. Poincaré inequality[11, 12, ?, 14]. Let Ω be a convex bounded region with diameter $\text{diam}(\Omega)$. Then, if f is in $W^{1,1}(\Omega)$ (it and its distributional gradient ∇f are integrable over Ω), and denoting the mean of f over Ω by $\langle f \rangle_{\Omega}$, the Poincaré inequality we need is

$$\int_{\Omega} |f - \langle f \rangle_{\Omega}| dx \leq \frac{\pi}{2} \text{diam}(\Omega) \int_{\Omega} |\nabla f| dx. \quad (5.3)$$

Applying this inequality to each cell of \mathfrak{P}^n and summing the results yields

$$\|\rho - \Pi_n \rho\|_1 \leq \frac{\pi\sqrt{3}}{2^{n+1}} \|\nabla \rho\|_1 \leq \frac{c}{2^n} F(\rho)^{1/2}, \quad \text{for } \rho \in \text{Dens}_{\mathcal{N}}. \quad (5.4)$$

Always, of course, $\|\rho - \Pi_n \rho\|_1 \leq 2\mathcal{N}$.

5.4.2. A dichotomy. In preparation for Thm. 5.1, we derive a useful inequality. For any measurable set B and $\rho, \rho' \in \text{Dens}_{\mathcal{N}}$,

$$\|\rho - \rho'\|_1 = \|(\rho - \rho')1_B\|_1 + \|(\rho - \rho')1_{B^c}\|_1,$$

Now we estimate the second term on the right-hand side.

$$\begin{aligned} \|(\rho - \rho')1_{B^c}\|_1 &\leq \int_{B^c} \rho dx + \int_{B^c} \rho' dx \leq 2 \int_{B^c} \rho dx + \left| \int_{B^c} (\rho' - \rho) dx \right| \\ &= 2 \int_{B^c} \rho dx + \left| \int_B (\rho' - \rho) dx \right| \leq 2 \int_{B^c} \rho dx + \|(\rho - \rho')1_B\|_1. \end{aligned}$$

The first inequality is a consequence of the fact that densities are non-negative. The equality in the second line follows because both densities ρ and ρ' have the same total integral (\mathcal{N}), and the final inequality is just pulling the absolute value inside the integral. Plugging this last result into the previous display,

$$\rho, \rho' \in \text{Dens}_{\mathcal{N}} \implies \|\rho - \rho'\|_1 \leq 2\|(\rho - \rho')1_B\|_1 + 2 \int_{B^c} \rho dx. \quad (5.5)$$

Theorem 5.1. Suppose $\rho_n \xrightarrow{wk-\mathfrak{P}} \rho$ in $\text{Dens}_{\mathcal{N}}$. Then, either $\rho_n \xrightarrow{L^1} \rho$ or $\liminf_n F^\circ(\rho_n) = +\infty$.

Proof. The equivalent nonstandard phrasing of the lemma is: if ρ' is $d_{\mathfrak{P}}$ -near ρ , then either it is L^1 -near ρ , or it has illimited internal energy. There is nothing to show unless $*F^\circ(\rho') \ll \infty$, so assume that is so.

The inequality (5.5) plays a key role, and a triangle inequality,

$$\|\rho - \rho'\|_1 \leq \|\rho - \Pi_n \rho\|_1 + \|\Pi_n(\rho - \rho')\|_1 + \|\rho' - \Pi_n \rho'\|_1, \quad (5.6)$$

will be used much as in Lemma 5.1. We proceed by showing that all three terms on the right-hand side are infinitesimal for some $N \in {}^*\mathbb{N}$. By definition of the weak- \mathfrak{P} topology, $\int_{C_n} |\Pi_n(\rho - \rho')| dx \simeq 0$ for every limited n [recall the definition of the C_n from (5.2)]. Robinson's lemma (§B.6) implies existence of an illimited N , fixed in the following discussion, such that that continues to hold for all $n \leq N$.

1st term of (5.6): ρ is standard and N illimited. By Lemma 5.1, $\|\rho - \Pi_N \rho\|_1 \simeq 0$.

2nd term: Apply (5.5) with C_N for B , and note that $\int_{C_N^c} \rho \simeq 0$ (ρ is standard). Thus, $\|\Pi_N(\rho - \rho')\|_1 \simeq 0$.

3rd term: ρ' has limited internal energy. Therefore, (5.4) implies $\|\rho' - \Pi_N \rho'\|_1 \simeq 0$. \square

Internal energy functionals are not upper semicontinuous because by putting in short-wavelength wiggles which are a negligible change in L^p (for any p) norm can drive the internal energy arbitrarily high. The inequality (5.1) and Thm. 5.4 show the flip-side of this phenomenon: under a bound on internal energy, we really only need to know a density down to a certain spatial resolution. Below that scale, it must be almost flat.

6. GROUND-STATE ENERGY, OR ‘ENERGY FLOOR’ (E)

At long last, we are ready to begin considering external one-body potentials. A one-body potential v induces the \mathcal{N} -body external potential $\underline{v}(\underline{x}) = \sum_{i=1}^{\mathcal{N}} v(x_i)$. The usual way of defining the ground-state energy is

$$E(v) = \inf \{ \mathcal{E}(\psi) + \langle \psi | \underline{v} \psi \rangle : \psi \in \text{State}^\circ \}.$$

Of course, the term $\langle \psi | \underline{v} \psi \rangle$ only depends on the state through its density, so the minimization can be done in two stages:

$$E(v) = \inf \{ F^\circ(\rho) + \langle v, \rho \rangle : \rho \in \text{Dens}_{\mathcal{N}} \}. \quad (6.1)$$

Here, we have introduced the notational abbreviation

$$\langle v, \rho \rangle := \int v \rho. \quad (6.2)$$

This is common in situations of pairing between elements of a Banach space and its dual, but we will use it even in cases which may yield $\langle v, \rho \rangle = +\infty$. Now, if the external potential v is bounded, then $\int v \rho$ is well-defined for all $\rho \in \text{Dens}_{\mathcal{N}}$ and the minimization in (6.1) makes sense. But if it is unbounded, we must proceed more cautiously.

6.0.1. on the term “ground-state energy”. The term “ground-state energy” for $E(v)$ is customary in the DFT literature. But, unless the infimum is actually a minimum, there is no genuine *ground state*. $v \equiv 0$ is the simplest example of a potential for which there is no minimum. A term, such as ‘energy floor’, which does not imply the existence of a ground state might therefore be more appropriate. Mostly, we shall avoid the issue by simply writing ‘ E ’.

6.0.2. Habitability and stability.

Definition 6.1. An external potential v is *habitable* if $E(v) < +\infty$ and *stable* if $E(v) > -\infty$.

Needless to say, we are interested in external potentials that are both habitable and stable, but, ensuring these two conditions involves very differing considerations.

6.0.3. some notation. $\langle v, \rho \rangle$ is not a function; it is the value of the function $x \mapsto \langle v, x \rangle$ at the argument ρ . To deal with this awkwardness, we write $\langle v, \cdot \rangle$ for the function $\rho \mapsto \langle v, \rho \rangle$. The notation

$$v^+ := v \vee 0, \quad v^- := (-v) \vee 0$$

for the positive and negative part of an external potential, such that $v = v^+ - v^-$ and $\langle v, \cdot \rangle = \langle v^+, \cdot \rangle - \langle v^-, \cdot \rangle$, will also be useful in this section.

6.1. Bounded external potentials. We begin by restricting our attention to external potentials in L^∞ , for which we use the notation ‘Bdd’.

If $w \in \text{Bdd}$, then $\langle w, \cdot \rangle$ is L^1 continuous, since L^∞ is the Banach space dual of L^1 . That is, $|\langle w, \rho \rangle - \langle w, \rho' \rangle| \leq \|w\|_\infty \|\rho - \rho'\|_1$. And since F is L^1 lsc, so too is $F + \langle w, \cdot \rangle$. However, Cor. 4.1 established that F is lsc with respect to much weaker topologies, so we should try to show the same for $F + \langle w, \cdot \rangle$. The full class of weak-but-not-leaky topologies is beyond our reach, but the weak- \mathfrak{P} topology can be handled.

Proposition 6.1. For $w \in \text{Bdd}$, $F + \langle w, \cdot \rangle : (\text{Dens}_{\mathcal{N}}, \text{wk-}\mathfrak{P}) \rightarrow \overline{\mathbb{R}}$ is lower semicontinuous.

Proof. Take ρ in $\text{Dens}_{\mathcal{N}}$ and ρ' wk- \mathfrak{P} -near ρ . According to Lemma 5.1, either $F(\rho') \simeq \infty$, or ρ' is L^1 -near ρ . In the former case, of course $F(\rho') \gtrsim F(\rho)$, because $|\int w \rho'| \leq \mathcal{N} \|w\|_\infty$ is bounded. In the latter, the conclusion follows by L^1 lsc of F and L^1 continuity of $\langle w, \cdot \rangle$. \square

6.1.1. *Fenchel conjugate and biconjugate*[15, 16, 17, 18]. We take a break to do some general geometric analysis. A proper (not identically $+\infty$) function $f : X \rightarrow \overline{\mathbb{R}}$ on a vector space X is given, along with a vector space Y of linear functionals $X \rightarrow \mathbb{R}$. The topology on X is that induced by Y , denoted $\sigma(X, Y)$. By definition, then, a set is a neighborhood of the origin if and only if it contains an intersection $\cap_n \{x : |\langle \xi_n, x \rangle| < \epsilon_n\}$, with $\xi_n \in Y$ and $\epsilon_n > 0$. In the $\sigma(X, Y)$ topology, a point x_0 is outside a closed convex set A if and only if there is $(\zeta, c) \in Y \times \mathbb{R}$ such that $\langle \zeta, x \rangle > c > \langle \zeta, x_0 \rangle$ for all $x \in A$, that is, x_0 can be strictly separated from A by a linear functional in Y . Here is an explanation for that. Forgetting about A 's convexity for the moment, there must be a collection $\{(\xi_i, c_i) : 1 \leq i \leq n\}$ and $\epsilon > 0$ such that $c_i = \xi_i(x_0)$, and for every $x \in A$, there is some i such that $|\langle \xi_i, x \rangle - c_i| > \epsilon$. Arrange the ξ_i 's into a map $\underline{\xi} = (\xi_1, \dots, \xi_n) : X \rightarrow \mathbb{R}^n$ into a finite-dimensional Euclidean space. Then, the image of A , $\underline{\xi}(A)$ is outside the ϵ ball around $\underline{c} = (c_1, \dots, c_n)$. But if A is convex, then so is its image in \mathbb{R}^n . We can apply *finite-dimensional* separation properties in \mathbb{R}^n to obtain a linear combination $\sum a_i \xi_i$ which achieves the desired separation.

The function defined on Y by

$$f^*(\zeta) := \sup \{ \langle \zeta, x \rangle - f(x) : x \in X \}, \quad (6.3)$$

is called, among other names, the *Fenchel conjugate*. Up to signs, the relationship between f and f^* is essentially the same as that between F and E . Note that

$$\forall x \in X. f(x) \geq c + \langle \zeta, x \rangle \Leftrightarrow -f^*(\zeta) \geq c.$$

Defining the half-space indexed by $(\zeta, c) \in Y \times \mathbb{R}$ by

$$H(\zeta, c) := \{(x, z) \in X \times \mathbb{R} : z \geq c + \langle \zeta, x \rangle\}, \quad (6.4)$$

we see that $\text{epi } f \subseteq H(\zeta, c)$ if and only if $-f^*(\zeta) \geq c$. (Recall that the epigraph of f is the set $\text{epi } f = \{(x, z) \in X \times \mathbb{R} : f(x) \leq z\}$ of points on or above the graph of f .)

The proposition we are aiming for is this: with \hat{f} the function satisfying $\text{epi } \hat{f} = \overline{\text{co}} \text{epi } f$,

$$\hat{f}(x) = \sup \{ \langle \zeta, x \rangle - f^*(\zeta) : \zeta \in Y \} = f^{**}(x). \quad (6.5)$$

An equivalent statement is that $\overline{\text{co}} \text{epi } f$ is equal to the intersection of all $H(\zeta, c)$ that contain $\text{epi } f$, and it is in this form that we prove it, by showing that any point *not* in $\text{epi } f$ falls outside some such half-space. We prove this assuming that $f \geq 0$ (adequate for our purposes), and later comment on the general case. Suppose $(x_0, z_0) \notin \overline{\text{co}} \text{epi } f$. Then, there is $(\zeta, \beta) \in Y \times \mathbb{R}$, $c \in \mathbb{R}$ and $\epsilon > 0$ such that

$$\forall (x, z) \in \text{epi } f. \langle \zeta, x \rangle + \beta z > c + \epsilon > c > \langle \zeta, x_0 \rangle + \beta z_0.$$

Since we can always increase z without leaving $\text{epi } f$, $\beta \geq 0$. If $x_0 \in \text{dom } f$, then $(x_0, z) \in \text{epi } f$ for large enough z , so $\beta > 0$ in that case. If $\beta > 0$, dividing through by β , we see that $\text{epi } f$ is in $H(-\zeta/\beta, c/\beta)$, but (x_0, z_0) is not. The other case is $\beta = 0$. Then,

$$c + \langle -\zeta, x \rangle < 0 < c + \langle -\zeta, x_0 \rangle, \quad \forall (x, z) \in \text{epi } f.$$

Thus, since $f \geq 0$, $f(x) > Mc + \langle -M\zeta, x \rangle$ for any $M > 0$. But, for large enough M , we will certainly have $z_0 < Mc + \langle -M\zeta, x_0 \rangle$. That completes the proof assuming $f \geq 0$. If that is not the case, since f is proper, there is some $x_0 \in \text{dom } f$. The β we get for this x_0 cannot be zero, as remarked earlier. Hence, $f + \langle \zeta/\beta, \cdot \rangle - c/\beta \geq 0$. If the proposition is true for this tilted and shifted function, it is also true for f .

6.1.2. *application to F* . Recall that $\text{epi } F = \overline{\text{co}} \text{epi } F^\circ$ (Thm. 4.3), where the closure is with respect to any weak-but-not-leaky topology on $\text{Dens}_{\mathcal{N}}$, such as L^1 or weak- \mathfrak{P} . Now, the weak- \mathfrak{P} topology is generated by the indicator functions of cells of \mathfrak{P} . Thus, to apply 6.1.1 we take Y to be

$$\text{Potl}^f(\mathfrak{P}) := \text{span} \{1_\Omega : \Omega \in \mathfrak{P}\}, \quad (6.6)$$

the vector space of functions spanned by the indicator functions of cells of \mathfrak{P} . A potential in $\text{Potl}^f(\mathfrak{P})$ is nonzero on only a finite number of cells. Then,

$$F(\rho) = \sup \left\{ E(v) - \langle v, \rho \rangle : v \in \text{Potl}^f(\mathfrak{P}) \right\}. \quad (6.7)$$

If we call a set of external potentials *F-sufficient* just in case the supremum over the functions $E(v) - \langle v, \cdot \rangle$ is F , then the display says that $\text{Potl}^F(\mathfrak{P})$ is F sufficient. Since always

$$E(v) \leq F(\rho) + \langle v, \rho \rangle, \quad (6.8)$$

any larger class of external potentials is also F -sufficient. We are not claiming that $\text{Potl}^F(\mathfrak{P})$ is a minimal F -sufficient class, but it is certainly pretty small.

Remark 6.1. Starting from L^1 lower-semicontinuity of F , one may similarly show the more-traditional, but weaker, result that L^∞ is F -sufficient. The proof uses the Hahn-Banach principle in the form of the assertion that the closures of a convex set relative to the initial and weak topology are the same.

6.2. $[0, +\infty]$ -valued external potentials. We now move from the relatively modest space Bdd of bounded external one-body potentials to a much larger class Potl^\uparrow . In contrast to the perturbations we consider later, it is not a vector space, but only a convex cone. This class consists of all measurable external potentials with values in the non-negative extended reals:

$$\text{Potl}^\uparrow = (\mathbb{R}^3 \rightarrow [0, +\infty]). \quad (6.9)$$

Why deal with such a large class of potentials? First, they have practical use in modeling various sorts of confining potentials, such as a harmonic oscillator potential or even an infinite cubic-well. Neither of these is in any L^p space, and a nice feature of our treatment is that such confining potentials are integrated into the *density* functional picture without any need to go back to talking explicitly about wavefunctions. Second, if we are going to deal with external potentials unbounded above, there is hardly any extra work to go all the way, for what we do. Third, and closely related to the second point, the differing *physical* significances of unboundedness above and below for an external potential is brought into the starkest possible relief.

A potential $w \in \text{Potl}^\uparrow$ can be approximated density-wise by *bounded* potentials. That is, with the truncations

$$(w \wedge n)(x) := (\text{if } w(x) < n \text{ then } w(x) \text{ else } n), \quad (6.10)$$

we have

$$\langle w \wedge n, \rho \rangle \nearrow \langle w, \rho \rangle, \text{ as } n \rightarrow \infty.$$

Thus,

$$F(\rho) + \langle w \wedge n, \rho \rangle \nearrow F(\rho) + \langle w, \rho \rangle.$$

The point is that if we can show that $\rho \mapsto F(\rho) + \langle w \wedge n, \rho \rangle$ is lower semicontinuous for each n , then so is the supremum (§2.3.1). But, we've already done that.

Theorem 6.1. *For $w \in \text{Potl}^\uparrow$, $F + \langle w, \cdot \rangle : (\text{Dens}_{\mathcal{N}}, \text{wk-}\mathfrak{P}) \rightarrow \overline{\mathbb{R}}$ is lower semicontinuous.*

Proof. By Prop. 6.1 and the preceding discussion. \square

6.2.1. sufficient condition for habitability of Potl^\uparrow external potentials.

Proposition 6.2. *$w \in \text{Potl}^\uparrow$ is habitable if $\int_U w < \infty$ for some open set $U \in \mathbb{R}^3$.*

Proof. Since U contains an open ball, one can easily find a smooth ρ with support in U , bounded by $\|\rho\|_\infty < \infty$. By inequality (4.21), $F(\rho) < \infty$, and $\langle w, \rho \rangle \leq \|\rho\|_\infty \int_U w < \infty$. \square

6.3. F -norm, and relatively F -small perturbations. An unbounded negative external potential poses a danger of *instability*. For such an external potential, we can find a sequence ρ_n of densities such that $\langle v, \rho_n \rangle \rightarrow -\infty$. Intuitively, this is achieved by squeezing the density more and more tightly into a bottomless potential well. To assure that the *total* energy cannot be driven to $-\infty$ in this way, we make sure that the kinetic energy cost of such a squeezing eventually dominates. This motivates the following definition.

Definition 6.2. The F -norm of a potential v is

$$\|v\|_F := \inf \{a \in (0, \infty] : \forall \rho \in \text{dom } F. \langle v, \rho \rangle \leq aF(\rho)\}. \quad (6.11)$$

A potential v is F -small if $\|v\|_F < 1$, and *relatively F -small* if it has a decomposition $v = v_1 + v_2$, with v_1 F -small and $\|v_2\|_\infty < \infty$. The set of all F -small potentials is denoted FSmall , and $\text{FSmall}^+ := \{v \geq 0 : v \in \text{FSmall}\}$.

Thus, the set of relatively F -small potentials is $\text{FSmall} + \text{Bdd}$. But, note that the decomposition of a potential into FSmall and Bdd pieces is not unique.

The appellation ‘ F -norm’ is justified because the norm conditions $\|v + w\|_F \leq \|v\|_F + \|w\|_F$, $\|av\|_F \leq |a|\|v\|_F$ (a real) and $\|v\|_F \Rightarrow v \equiv 0$ are satisfied. These imply that $\|\cdot\|_F$ is convex. A subset S of a vector space is called *balanced* if $x \in S$ implies $-x \in S$.

Lemma 6.1. FSmall is convex and balanced.

With that, we are ready to define our big space of potentials.

Definition 6.3.

$$\begin{aligned} \text{Potl} &:= \text{Potl}^\uparrow + \text{Bdd} + \text{FSmall} \\ &= \text{Potl}^\uparrow + \text{Const} - \text{FSmall}^+. \end{aligned} \quad (6.12)$$

Note that neither of the exhibited decompositions is unique, but the second is closer to being so.

Theorem 6.2. Every external potential in Potl is stable.

Proof. Let v be decomposed as $v = w + u - u' \in \text{Potl}^\uparrow + \text{Const} - \text{FSmall}^+$. Then, since $w + u$ is bounded below, so is $\langle w + u, \rho \rangle$ independently of $\rho \in \text{Dens}_{\mathcal{N}}$. Also,

$$F(\rho) - \langle u', \rho \rangle \geq (1 - \|u'\|_F)F(\rho) > 0.$$

□

6.4. $L^{3/2} + L^\infty$. Conceptually, the F -norm is the right thing. Unfortunately, its computation poses significant problems. However, if we have a tractable upper bound, we will have sufficient conditions for an external potential to be stable. This is where Lieb’s space of potentials $L^\infty + L^{3/2} = \text{Bdd} + L^{3/2}$ comes in. Recall the inequality (4.22): $\|\rho\|_3 \leq cF(\rho)$. This does not say that an $L^{3/2}$ potential is F -small. But, if we chop off a big enough bounded part, the remainder will be so.

Lemma 6.2. $L^{3/2} \subseteq \text{Bdd} + \text{FSmall}$.

Proof. v^+ and v^- can be treated separately, so, without loss, assume $v \geq 0$. Then, $v \wedge n \rightarrow v$ in $L^{3/2}$ norm, so there is an n such that $v' = v - (v \wedge n)$ satisfies $\|v'\|_{3/2} < a/c$, with $a < 1$. But then, $|\langle v', \rho \rangle| \leq \|v'\|_{3/2} \|\rho\|_3 < aF(\rho)$ for every $\rho \in \text{Dens}_{\mathcal{N}}$. (The first inequality is Hölder’s). And since $|v \wedge n| \leq n$, $v = (v \wedge n) + v'$ is the required decomposition. □

6.5. **Concavity and continuity.** This section treats concavity and continuity properties of E .

6.5.1. *concavity.* A function f is *concave* if $-f$ is convex. So E is concave if its domain is convex and for all v and w in its domain,

$$\forall \alpha \in (0, 1). E(\alpha v + (1 - \alpha)w) \geq \alpha E(v) + (1 - \alpha)E(w).$$

The formula (6.1) expresses E as the infimum of a collection of functions $F(\rho) + \langle \cdot, \rho \rangle$ indexed by densities. If we phrase this in terms of subgraphs (sub f , the *subgraph* of f , is the region on or below the graph of f),

$$\text{sub } E = \bigcap_{\rho \in \text{Dens}_{\mathcal{N}}} \text{sub } (F(\rho) + \langle \cdot, \rho \rangle).$$

If every function on the right is concave, then their subgraphs are convex, hence the intersection of their subgraphs is convex, and therefore E is also concave. We therefore need to verify the condition that for each $\rho \in \text{Dens}_{\mathcal{N}}$, $v, w \in \text{Potl}$ and $0 < \alpha < 1$,

$$F(\rho) + \langle \alpha v + (1 - \alpha)w, \rho \rangle \geq \alpha [F(\rho) + \langle v, \rho \rangle] + (1 - \alpha) [F(\rho) + \langle w, \rho \rangle].$$

Decompose $v = v' + v''$ with $v' \in \text{Potl}^\uparrow + \text{Const}$ and $v'' \in \text{FSmall}$, and similarly, $w = w' + w''$. Now,

$$\langle \alpha v' + (1 - \alpha)w', \rho \rangle = \alpha \langle v', \rho \rangle + (1 - \alpha) \langle w', \rho \rangle,$$

so we need show only

$$F(\rho) - \langle \alpha v'' + (1 - \alpha)w'', \rho \rangle \geq \alpha [F(\rho) - \langle v'', \rho \rangle] + (1 - \alpha) [F(\rho) - \langle w'', \rho \rangle].$$

But, since v'' and w'' are F -small, if either potential term on the right-hand side is infinite, then $F(\rho) = +\infty$, so both sides are infinite. Thus, the following is now established.

Proposition 6.3. $E : \text{Potl} \rightarrow \overline{\mathbb{R}}$ is concave.

6.5.2. *A locally lower-bounded convex function is locally Lipschitz.* Convexity of a function is an essentially two-dimensional affair, and it allows us to usefully apply elementary geometry to function spaces. A simple drawing shows that slopes of secant lines to a convex function $f : \mathbb{R} \rightarrow \mathbb{R}$ are non-decreasing: $z > y \geq x > w$ implies $\frac{f(z) - f(y)}{z - y} \geq \frac{f(x) - f(w)}{x - w}$. From this observation, we immediately deduce the following.

Lemma 6.3. *Let $f : [-a, a] \rightarrow \mathbb{R}$ with $f(0) = 0$ be convex and bounded above by M . Then f is bounded below by $-M$. Furthermore, it satisfies $|f(x)| \leq (M/a)|x|$ (locally Lipschitz at 0 with Lipschitz constant M/a).*

We have phrased this lemma in terms of a convex function. Of course, it has an analog for a concave function: locally upper-bounded implies locally Lipschitz.

6.5.3. *“Perturbative” continuity of E .* For $w \in \text{Potl}$, we think of $w + v$ for $v \in \text{Bdd} + \text{FSmall}$ as a perturbation of the “background” w and write

$$E_w(v) = E(w + v). \quad (6.13)$$

For fixed background w , $E_w(v)$ is locally Lipschitz as a function of v .

Theorem 6.3. *If $w \in \text{Potl}$ is habitable, then $E_w(\cdot) : \text{Bdd} + \text{FSmall} \rightarrow \mathbb{R}$ is locally Lipschitz at zero.*

Proof. Given $\epsilon > 0$, choose ρ_0 such that

$$F(\rho_0) + \langle w, \rho_0 \rangle < E(w) + \epsilon. \quad (6.14)$$

Then, for $v \in \text{Bdd}$ and $v' \in \text{FSmall}$,

$$E_w(v + v') \leq F(\rho_0) + \langle w, \rho_0 \rangle + \langle v + v', \rho_0 \rangle < E(w) + \epsilon + \mathcal{N}\|v\|_\infty + F(\rho_0)\|v\|_F.$$

Thus, E_w is locally bounded above, and Lemma 6.3 finishes the proof. \square

This brings us to the end of our investigation of the functionals F and E . In 6.1.2 we just scratched the surface of the relationship between them, and have not even touched the vexed v -representability problem of finding a potential with given ground-state density. Nevertheless, it is time to quit, for now.

APPENDIX A. SELECTIVE FUNCTIONAL ANALYSIS REFRESHER

This appendix contains a refresher on some basic notions of functional analysis. The reader for whom this is completely new is likely to feel the need for additional resources.

A.1. Lebesgue (L^p) and Sobolev (H^1 , $W^{1,1}$) spaces. We recall here the definitions of L^p norms. For any measurable function $f : \mathbb{R}^n \rightarrow \mathbb{R}$,

$$\begin{aligned} \|f\|_{L^p}^p &= \|f\|_p^p := \int |f(x)|^p dx, \quad 0 < p < \infty \\ \|f\|_{L^\infty} &= \|f\|_\infty := \inf\{a \geq 0 : \mu_{\text{Leb}}(|f| \leq a) > 0\}. \end{aligned} \quad (\text{A.1})$$

In the second line, μ_{Leb} denotes Lebesgue measure. For a continuous function, $\|\cdot\|_\infty$ is just the supremum of the absolute value. The complicated business with μ_{Leb} is needed so that changing a function on a set of measure zero does not change its norm. The space $L^p(\mathbb{R}^n)$ consists of the

functions with finite L^p norm. Actually, this is a white lie. None of the $\|\cdot\|_p$ can distinguish functions which differ on sets of measure zero, so we really should say that the elements are *equivalence classes of functions* rather than functions, *tout court*. But nobody does.

Sobolev spaces involve also norms of derivatives. These are really distributional derivatives, and for our purposes, it is simplest to think in terms of Fourier transform. With

$$\hat{f}(q) = \lim_{R \rightarrow \infty} \int_{|x| \leq R} e^{iq \cdot x} f(x) \frac{dx}{(2\pi)^{n/2}}$$

the Fourier transform of f , the H^1 norm of f is

$$\|f\|_{H^1}^2 := \int (1 + |q|^2) |\hat{f}(q)|^2 dq.$$

If f is actually differentiable, then this is the same as

$$\|f\|_{H^1}^2 := \int (|f|^2 + |\nabla f|^2) dx = \|f\|_2^2 + \|\nabla f\|_2^2,$$

where the derivative is understood in the classical sense. Otherwise ∇f is a distributional derivative. The other Sobolev space used in these notes is $W^{1,1}$:

$$\|f\|_{W^{1,1}} := \int (|f| + |\nabla f|) dx.$$

The superscripts in ‘ $W^{1,1}$ ’ indicate maximum derivative order and the power to which things are raised in the integral. H^1 is also called $W^{1,2}$. The reader will see that these are just the beginning of a family of function spaces. Many sources are available for more information, for example [19, 20, 21, 11, 22].

A.2. the concept of ‘topology’. ‘Topology’ is a badly overloaded word, largely due to a tendency to deploy it in contexts (turning coffee cups into a doughnuts) where a more precise term would be better. As we use it here, a *topology* is essentially a notion of convergence. In a metric space, a sequence (x_n) converges to x if the distances $d(x_n, x) \rightarrow 0$ in \mathbb{R} . What is involved here are certain standards of closeness, technically called *neighborhoods*. The ball $B_x(a)$ of radius a about x is one neighborhood of x and the ball $B_x(b)$ is another. If $a > b$, the second standard of closeness is strictly more demanding than the first; $B_x(b) \subset B_x(a)$. We *could* consider “lopsided” neighborhoods which do not nest. The point for a metric space is that we do not need such things; the notion of convergence is completely and cleanly captured by the centered balls. But, such a simple structure is not always obtainable. Consider the space of bounded functions $\mathbb{R} \rightarrow \mathbb{R}$ where we say that $f_n \rightarrow f$ if and only if $f_n(x) \rightarrow f(x)$ for each x . Example: $f_n(x) = (\text{if } x > n \text{ then } 1 \text{ else } 0)$ converges to the function 0, which is identically equal to zero. $\{g : |g(a)| < \epsilon\}$ is one neighborhood of 0 and $\{g : |g(b)| < \epsilon\}$ is another. The two are not comparable, and we cannot rejigger to eliminate this complication. But, there is a third neighborhood, $\{g : |g(x)| < \epsilon \text{ for } x \in \{a, b\}\}$ which is contained in both. This illustrates the general pattern. One way to specify the topology of a topological space is to give, for each point x , a *neighborhood base* \mathcal{B}_x such that if $A, B \in \mathcal{B}_x$, then there is $C \in \mathcal{B}_x$ satisfying $C \subseteq A, C \subseteq B$. Often, given such a system of neighborhood bases, we can find a metric which has the same concept of convergence. This is done in §5.1.1, for example. In fact, no non-metrizable topology is essential anywhere in these notes, so the reader loses nothing by reading every occurrence of “topology” in the text as “notion of convergence associated to some metric”.

APPENDIX B. INFINITESIMALS

B.1. seriously. Infinitesimals have a long history in mathematics. Leibniz based his version of the calculus on them, and Euler made masterful use of infinitesimals and their reciprocals, infinitely large quantities. But the logical status of old-style infinitesimals was ever shaky, and they were banished from mathematical discourse during the 19th century, replaced by the ϵ - δ method. Eventually, infinitesimals were revived in a rigorous form, primarily through the work of Abraham Robinson in the 1960’s, giving us modern *infinitesimal analysis*, or *nonstandard analysis*[23]. Although there

are alternate axiomatic treatments[24, 25, 26], we follow the model theoretic approach pioneered by Robinson[23, 27, 28, 29, 30, 31, 32, 33]. The idea is to extend a conventional mathematical universe \mathbb{U} to a nonstandard universe $^*\mathbb{U}$ which has all the same *formal* properties as \mathbb{U} does, and yet is chock full of ideal elements. That an element is ideal, for example infinitesimal, is an *external* judgement not available to the *internal* object language. The interplay of internal and external is responsible for the power of infinitesimal methods. We give here an informal exposition. The previously-cited references give details of the customary construction of nonstandard universes by means of ultraproducts over set-theoretic superstructures. Relieved of the burden of proving anything, we describe matters instead in a type-theoretic way inspired by modern functional programming languages such as ML and Haskell. A good strategy for the reader might be to go through just B.2–B.4, and then return when necessary. The only nonstandard argument which is not elementary is contained in the proof of Lemma 3.1.

B.2. base datatypes. We begin with the base datatypes of Booleans $\text{Bool} = \{\text{T}, \text{F}\}$, natural numbers \mathbb{N} and real numbers \mathbb{R} . These types, and their elements, are denizens of the mathematical universe \mathbb{U} . The nonstandard universe $^*\mathbb{U}$ has $^*\text{Bool}$, which is actually identical to Bool , the hypernaturals $^*\mathbb{N}$ extending \mathbb{N} and the hyperreals $^*\mathbb{R}$ extending the reals. $^*\mathbb{N}$ and $^*\mathbb{R}$ have all the order and arithmetic properties of \mathbb{N} and \mathbb{R} . But these sets have additional elements; the new elements in $^*\mathbb{N}$ are larger than all naturals and some of the hyperreals are infinitesimal. With the understanding of implicit coercion from \mathbb{N} to \mathbb{R} and from $^*\mathbb{N}$ to $^*\mathbb{R}$, we have the following grab-bag of terminology and ideas. The essential image is of a number line with a “halo” of nonstandard numbers around every standard real and then extended on the far left and far right with “infinitely large” numbers.

Definition B.1. Let $x, y \in ^*\mathbb{R}$.

- a. x is *standard* if $x \in \mathbb{R}$. Otherwise, x is *nonstandard*.
- b. x is *limited* if $|x| < n$ for some $n \in \mathbb{N}$; otherwise x is *illimited*. There is no smallest illimited hypernatural or hyperreal.
- c. x is *infinitesimal* if $|x| < 1/n$ for all $n \in \mathbb{N}$; otherwise x is *appreciable*.
- d. If $x - y$ is infinitesimal, then x and y are *infinitely close*, or simply *near*. Notation: $x \simeq y$.
- e. If x is limited, then $x \simeq y$ for a unique $y \in \mathbb{R}$. y is the *standard part* of x , denoted by ${}^\circ x$ or $\text{st } x$.
- f. If x is standard, the collection of all hyperreals near x makes up the *halo* (or *monad*) of x . Halos of distinct reals are disjoint.
- g. $x \gg y$ means $x > y$ but $x \neq y$. $x \ll \infty$ is a synonym for “ x is limited”.

Any formal statement S we can make about Bool , \mathbb{N} , and \mathbb{R} using Boolean operations (conjunction \wedge , disjunction \vee and negation \neg), natural and real constants and arithmetic operations $(0, 1, +, \cdot, <, | \cdot |)$ and bounded quantifiers (\forall and \exists), is turned into a statement *S about Bool , $^*\mathbb{N}$ and $^*\mathbb{R}$ by putting stars on the constants and parameters. Leibniz’ Principle, which we require, says that S and *S have the same truth value. For example, every natural is either even or odd, hence the same is true of a hypernatural, even if it is illimited. Every hypernatural has a successor. There is a third hyperreal strictly between any two distinct hyperreals. There is no real number squaring to -1 , hence neither is there such a hyperreal. Since we *identify* reals, naturals and Booleans with their $*$ -transforms, “ $^*(-1)$ ” is simply “ -1 ”.

B.3. limits and continuity. With individual functions added to our object language, we can discuss nonstandard characterizations of sequence limits and function continuity. These are bread-and-butter applications of infinitesimal methods. A real sequence (x_n) is really a function $\mathbb{N} \rightarrow \mathbb{R}$, and has an extension $(^*x_n)$ with n running over $^*\mathbb{N}$. We want to show the equivalence of “ $x_n \rightarrow a$ ” and “ $^*x_n \simeq a$ for $n \simeq \infty$ ”.

The formalization of $x_n \rightarrow a$ is $(\mathbb{R}_{>0}$ abbreviates the reals greater than zero) $\forall \epsilon \in \mathbb{R}_{>0} \exists n \in \mathbb{N}. m > n \Rightarrow |x_m - a| < \epsilon$. Actually, we do not want to put ϵ and n in the formal statement, but keep them at the metalinguistic level. Then we have simply,

$$\forall m > n(\epsilon) \Rightarrow |x_m - a| < \epsilon,$$

which is true in \mathbb{U} . Transfer says (recall, n is the same as *n and similarly for ϵ)

$$\forall m > n(\epsilon) \Rightarrow |{}^*x_m - a| < \epsilon.$$

Now, if $m \simeq \infty$, then certainly $m > n(\epsilon)$. And that's true no matter how small ϵ is, so $x_m \simeq a$. For the other direction, suppose we know that $m \simeq \infty$ implies $x_m \simeq a$. Then, given standard $\epsilon > 0$, the statement

$$\exists n \in {}^*\mathbb{N}. m > n \Rightarrow |{}^*x_m - a| < \epsilon$$

is true in ${}^*\mathbb{U}$; any illimited n will serve. Backward transfer gives us

$$\exists n \in \mathbb{N}. m > n \Rightarrow |x_m - a| < \epsilon.$$

Since this procedure is valid for any $\epsilon \in \mathbb{R}_{>0}$, we conclude that $x_n \rightarrow a$.

Continuity of a function $f : \mathbb{R} \rightarrow \mathbb{R}$ at a has an infinitesimal characterization with a very similar flavor. It is “ $x \simeq a$ implies ${}^*f(x) \simeq {}^*f(a)$ ”, and is demonstrated by a very similar technique. Ordinarily, we might drop the stars on “ $f(x)$ ” and “ $f(a)$ ”, as there is neither ambiguity nor possibility for confusion (Not even externally does “ $a \simeq b$ ” make sense for $a, b \in \mathbb{R}$.)

B.4. metric, function, and topological spaces. Any metric space X is extended in much the same way as \mathbb{R} . For example, consider the space ℓ^2 of square summable sequences $x : \mathbb{N} \rightarrow \mathbb{R}$, with the norm $\|x\|_2^2 = \sum_n |x_n|^2$. Then ${}^*\ell^2$ consists of sequences ${}^*\mathbb{N} \rightarrow {}^*\mathbb{R}$. $x \in \ell^2$ has an extension ${}^*x \in {}^*\ell^2$; it is a standard element. Now, since $\sum |x_n|^2$ converges, for any standard $\epsilon > 0$, there is $n(\epsilon)$ such that $\sum_{m \geq n(\epsilon)} |x_m|^2 < \epsilon$. So using Transfer, we conclude that $\sum_{m \geq N} |x_m|^2 \simeq 0$ for any illimited N . This is important to the nature of nearstandard elements of ${}^*\ell^2$. If $y \in {}^*\ell^2$ has a standard part, then $\sum_{n \in {}^*\mathbb{N}} |(\text{st } y)_n - y_n|^2 \simeq 0$. This implies that $(\text{st } y)_n \simeq {}^\circ(y_n)$ for limited n , as well as $\sum_{n \in \mathbb{N}} |{}^\circ(y_n)|^2 \ll \infty$ and $\sum_{n \geq N} |y_n|^2 \simeq 0$ for any illimited N . We see, then, that there are basically two ways $y \in {}^*\ell^2$ can fail to be nearstandard (not nearstandard is also called *remote*). First, it could have an illimited norm, $\|y\| \approx \infty$. Second, it could have appreciable weight at infinity. The first has an analog for ${}^*\mathbb{R}$, but the second does not. This is the nonstandard manifestation of the fact that bounded subsets of \mathbb{R} are relatively compact, but bounded subsets of ℓ^2 are not. The element $y_n = \delta_{nN}$ (n illimited) is a remote element with norm one.

If d is a metric for X in \mathbb{U} , then *d is a metric for *X . Each standard point $x \in {}^*X$ is surrounded by a halo (or ‘monad’) of nonstandard points at infinitesimal distance from it. If $d(y, {}^*x) \simeq 0$, then x is the *standard part* of y , written either ${}^\circ y$ or $\text{st } y$, as convenient. *X may well have points which are not near any standard point. These are *remote*. Indeed, if there are points in X arbitrarily far from x ($\forall n \in \mathbb{N} \exists y \in X. d(y, x) > n$), then there are points in *X an illimited distance from *x . These are necessarily remote, since any putative standard part must be a finite distance from x . A function $f : X \rightarrow \mathbb{R}$ is continuous at $x \in X$ if and only if *f maps the halo of x into the halo of $f(x)$. All these points are exemplified by the space ℓ^2 of the previous paragraph.

We touch briefly on the generalization to an arbitrary topological space in \mathbb{U} . There is only one essential difference. Without a metric, there is no numerical measure of “infinitely close”. Instead, the halo of a standard point x must be defined as $\cap {}^*U$, where the intersection runs over all the standard open neighborhoods of x . One can check that for a metric space this matches the definition we have already given.

B.5. *-transform and Leibniz’ Principle. To do much interesting mathematics, we need higher-order types, such as power sets and function spaces, that are so far lacking from the language. If A and B are datatypes in \mathbb{U} , then so is the product type $A \times B$ consisting of all ordered pairs (x, y) for $x \in A$ and $y \in B$, the sum (disjoint union) type $A + B$, and the type $[A \rightarrow B]$ of all functions from A to B . With these datatype constructors, along with function abstraction and application, we can fill out the universe \mathbb{U} and extend the * -transform to a map $\mathbb{U} \xrightarrow{*} {}^*\mathbb{U}$ obeying the Transfer Principle.

- a. The range of * is the *standard* objects. These are raw materials from which the other, *internal*, objects of ${}^*\mathbb{U}$ are obtained by ordinary mathematical operations (*de novo* function space construction is not ordinary).

- b. Individual Booleans, naturals and reals are urelements with no internal structure; we identify them with their $*$ -transforms.
- c. If $x \in A$, then $*x \in *A$. But if A is infinite, $*A$ contains nonstandard members as well (example: illimited hypernaturals).
- d. $*$ respects product and sum construction, $*(A \times B) = *A \times *B$, $*(A + B) = *A + *B$. , tupling, $*(x, y) = (*x, *y)$, and function application, $*(f(x)) = *f(*x)$.
- e. Only the function space constructor has an abnormal interpretation in $*\mathbb{U}$. $*[A \rightarrow B]$ is only a subset of $[*A \rightarrow *B]$, which we write as $[*A \dashrightarrow *B]$. Members of $[*A \dashrightarrow *B]$ are *internal* functions.
- f. The power set constructor, $\mathcal{P}(\cdot)$, is basically syntactic sugar for $[\cdot \rightarrow \text{Bool}]$. Other set operations are likewise abbreviations, which we use freely.
- g. $*\mathbb{U}$ consists of *internal* objects, anything we can describe using the resources of the formal language and beginning with standard objects. In particular, a standard set is internal, any *member* of a (member of a \dots of a) standard set is internal. If $c \in *\mathbb{R}$ is infinitesimal, then it is not standard (not of the form $*b$), but it is internal, and so is the subset $[0, c]$ of $*\mathbb{R}$.
- h. If C and D are internal subsets of $*A$ and $*B$, respectively, $[C \dashrightarrow D]$ is obtained in the normal way by selection and restriction. Also, the internal function space constructor respects currying: if A, B and C are internal, then $A \dashrightarrow [B \dashrightarrow C] = (A \times B) \dashrightarrow C$.
- i. The symbol ' \simeq ', the predicate 'limited', the standard part operation and the other elements of Definition B.1 are *not* part of the object language. Most especially, neither $\{x \in *\mathbb{N} : x \ll \infty\}$ nor $\{x \in *\mathbb{R} : x \simeq 0\}$ is internal. They are *external*. Because we can formally express the least-number principle (any nonempty subset of \mathbb{N} has a least element), Transfer implies that it holds for *internal* subsets of $*\mathbb{N}$. But, there is no smallest illimited hypernatural. Consistency thus demands that these things be external sets.

B.6. overspill. 'Limited' and 'infinitesimal' are the most important external concepts. We illustrate a typical pattern of use. Suppose a sequence $x : *\mathbb{N} \rightarrow *\mathbb{R}$ has the property that $nx_n < 1$ for every limited n . Then, the set $B = \{m \in *\mathbb{N} : \forall n < m, nx_n < 1\}$ is evidently internal. Therefore, B cannot consist of only the limited hypernaturals, since that is an external set. There must be some illimited m in B . Thus the property $nx_n < 1$ *spills over* into the illimited hypernaturals. Similarly, if a formally expressible property holds for all appreciable hyperreals, it must hold for some infinitesimal ones. Now suppose an internal sequence (x_n) satisfies $x_n \simeq 0$ for all limited n . Can we use overspill? Since " \simeq " is external, it is not evidently applicable, but there is a clever trick discovered by Robinson. The given condition implies $nx_n < 1$ for limited n . And for illimited n , $nx_n < 1$ implies $x_n \simeq 0$. Thus the property $x_n \simeq 0$ actually does extend into the illimited hypernaturals. This result, known as Robinson's Lemma, is used in 5.4.2.

REFERENCES

- [1] Lieb E H 1983 *Int J Quantum Chem* **24** 243
- [2] Parr R and Yang W 1989 *Density-Functional Theory of Atoms and Molecules* (Cambridge: Clarendon)
- [3] Dreizler R M and Gross E K U 1990 *Density Functional Theory: an approach to the quantum many-body problem* (Berlin: Springer-Verlag)
- [4] Eschrig H 1996 *The Fundamentals of Density Functional Theory* (Stuttgart, Leipzig: Teubner)
- [5] Martin R M 2004 *Electronic Structure: Basic Theory and Practical Methods* (Cambridge, New York: Cambridge University Press)
- [6] Capelle K 2006 *Braz. J. Phys.* **36** 1318–1343 arXiv:cond-mat/0211443
- [7] van Leeuwen R 2003 density functional approach to the many-body problem: key concepts and exact functionals *Advances in Quantum Chemistry* vol 43 ed R S J and E B (Amsterdam: Elsevier) pp 25–94
- [8] Levy M 1982 *Phys. Rev. A* **26** 1200–1208
- [9] Harriman J E 1981 *Phys. Rev. A* **24** 680–682
- [10] Zumbach G 1985 *Phys. Rev. A* **31** 1922–1925
- [11] Adams R A 1975 *Sobolev Spaces* (New York: Academic Press)
- [12] Attouch H, Buttazzo B and Michaille G 2006 *Variational Analysis in Sobolev and BV Spaces* (Philadelphia: SIAM)
- [13] Acosta G and Durán R G 2003 *Proc. Am. Math. Soc.* **132** 195–202
- [14] Beberndorf M 2003 *Zeit. für Analysis und ihre Anwendungen* **22** 751–756

- [15] Ekeland I and Témam R 1976 *Convex Analysis and Variational Problems* (Amsterdam: North-Holland) reprinted 1999 (SIAM, Philadelphia)
- [16] Aubin J P and Ekeland I 1984 *Applied Nonlinear Functional Analysis* (New York: Wiley) reprinted (Dover, Mineola, NY, 2006)
- [17] Borwein J M and Zhu Q J 2005 *Techniques of Variational Analysis* (New York: Springer-Verlag)
- [18] Phelps R R 1988 *Convex functions, monotone operators and differentiability* 2nd ed (*Lecture Notes in Mathematics* vol 1364) (Berlin, New York: Springer-Verlag)
- [19] Taylor M 1996 *Partial Differential Equations: Basic Theory* (New York: Springer-Verlag)
- [20] Trèves F 2006 *Basic linear partial differential equations* (Mineola, NY: Dover Publications Inc.) ISBN 0-486-45346-4 reprint of the 1975 original
- [21] Folland G B 1995 *Introduction to partial differential equations* 2nd ed (Princeton, NJ: Princeton University Press) ISBN 0-691-04361-2
- [22] Lieb E H and Loss M 1997 *Analysis* (Providence, R.I.: American Mathematical Society)
- [23] Robinson A 1966 *Non-standard Analysis* (Amsterdam: North-Holland) reissued by Princeton University Press, 1996
- [24] Nelson E 1977 *Bull. Amer. Math. Soc.* **83** 1165–1198 ISSN 0002-9904
- [25] Nelson E 1987 *Radically elementary probability theory* (*Annals of Mathematics Studies* vol 117) (Princeton, NJ: Princeton University Press) ISBN 0-691-08473-4; 0-691-08474-2
- [26] Kanovei V and Reeken M 2004 *Nonstandard analysis, axiomatically* Springer Monographs in Mathematics (Berlin: Springer-Verlag) ISBN 3-540-22243-X
- [27] Lindström T 1988 *Nonstandard Analysis* ed Cutland N (Cambridge, New York: Cambridge University Press) pp 1–105
- [28] Goldblatt R 1998 *Lectures on the Hyperreals: an introduction to nonstandard analysis* (New York: Springer)
- [29] Loeb P A and Hurd A E 1985 *An introduction to nonstandard real analysis* (Orlando: Academic Press)
- [30] Loeb P A 2000 *Nonstandard Analysis for the Working Mathematician* ed Loeb P A and Wolff M P A (Dordrecht, Boston: Kluwer Academic Publishers)
- [31] Davis M 1977 *Applied Nonstandard Analysis* (New York: Wiley)
- [32] Alberverio S, Fenstad J, Hoegh-Krohn R and Lindström T 1986 *Nonstandard methods in Stochastic Analysis and Mathematical Physics* (Orlando: Academic Press) reprinted by Dover, 2009
- [33] Stroyan K D and Luxemburg W A J 1976 *Introduction to the Theory of Infinitesimals* (New York: Academic Press)